

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: August 8, 2003, 14:36:31 ; Search time 2678 Seconds  
(without alignments)  
12143.162 Million cell updates/sec

Title: US-09-806-536A-29  
Perfect score: 1338  
Sequence: 1 ggtgagcgcagtctgtccga.....ggggcacgacataggggt 1338

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

- EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_esthm:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_htc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_htc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rtd:\*  
26: em\_gss\_pbg:\*  
27: em\_gss\_vrl:\*  
28: gb\_gss1:\*  
29: gb\_gss2:\*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | ID | Description |
|------------|-------|---------------|--------|----|-------------|
| 1          | 980.2 | 73.3          | 1201   | 13 | BX352919    |
| 2          | 889.2 | 66.5          | 1201   | 9  | AL519338    |
| 3          | 884   | 66.1          | 1086   | 12 | BW548059    |
| 4          | 880.6 | 65.8          | 977    | 9  | AL526237    |

|    |       |      |      |    |          |
|----|-------|------|------|----|----------|
| 5  | 832.2 | 62.2 | 1080 | 12 | BM923064 |
| 6  | 794.6 | 59.4 | 873  | 9  | AL525111 |
| 7  | 786.4 | 58.8 | 873  | 12 | BI837148 |
| 8  | 763.8 | 57.1 | 773  | 12 | BI827333 |
| 9  | 753   | 56.3 | 856  | 12 | BM014974 |
| 10 | 737.6 | 55.1 | 794  | 9  | AU141625 |
| 11 | 737.2 | 55.1 | 964  | 13 | BX370424 |
| 12 | 704.8 | 52.7 | 884  | 13 | BU181305 |
| 13 | 701.2 | 52.4 | 1145 | 12 | BM473687 |
| 14 | 697.4 | 52.1 | 932  | 13 | BQ436653 |
| 15 | 688.2 | 51.4 | 734  | 12 | BI253559 |
| 16 | 683.2 | 51.1 | 770  | 9  | AL522085 |
| 17 | 660   | 49.3 | 741  | 12 | BM010281 |
| 18 | 627.6 | 46.9 | 773  | 12 | BI818555 |
| 19 | 622.4 | 46.5 | 664  | 9  | AI963473 |
| 20 | 620.4 | 46.4 | 679  | 10 | BG393676 |
| 21 | 617   | 46.1 | 895  | 10 | BF797243 |
| 22 | 614.4 | 45.9 | 1036 | 12 | BM45513  |
| 23 | 601.4 | 44.9 | 890  | 13 | BU172522 |
| 24 | 589.8 | 44.1 | 1024 | 12 | BM465183 |
| 25 | 580.4 | 43.4 | 590  | 10 | BG282890 |
| 26 | 569.2 | 42.5 | 1271 | 11 | AK002912 |
| 27 | 568.8 | 42.4 | 1201 | 9  | AL556719 |
| 28 | 567.8 | 42.4 | 823  | 12 | BM019985 |
| 29 | 558.4 | 41.7 | 560  | 9  | AU280528 |
| 30 | 557.4 | 41.7 | 595  | 10 | BE621964 |
| 31 | 553.6 | 41.4 | 902  | 14 | CB182842 |
| 32 | 552.8 | 41.3 | 1132 | 12 | BM424026 |
| 33 | 539.6 | 40.3 | 996  | 14 | BY703409 |
| 34 | 535.8 | 40.0 | 1265 | 11 | AK013214 |
| 35 | 535   | 40.0 | 803  | 13 | BQ571032 |
| 36 | 532   | 39.8 | 744  | 10 | BE267359 |
| 37 | 532   | 39.8 | 902  | 13 | BQ964436 |
| 38 | 527.6 | 39.4 | 788  | 12 | BI254417 |
| 39 | 525.6 | 39.3 | 874  | 12 | BI415573 |
| 40 | 520.8 | 38.9 | 1348 | 11 | AK004291 |
| 41 | 520.6 | 38.9 | 737  | 12 | BM019007 |
| 42 | 517.4 | 38.7 | 670  | 10 | BE563597 |
| 43 | 510.6 | 38.2 | 869  | 12 | BI329060 |
| 44 | 505.2 | 37.8 | 602  | 9  | AL706316 |
| 45 | 504.4 | 37.7 | 1178 | 12 | BM461077 |

ALIGNMENTS

RESULT 1  
BX352919  
LOCUS  
DEFINITION BX352919 Homo sapiens NEUROBLASTOMA COT 10-NORMALIZED Homo sapiens  
CDNA clone CS0DB006YC20 5-PRIME, mRNA sequence.  
ACCESSION BX352919  
VERSION BX352919.1 GI:30347710  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE  
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished  
COMMENT Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of  
Invitrogen. This sequence belongs to sequence cluster 4390.r For  
more information about this cluster, see  
http://www.genoscope.cns.fr/  
cgi-bin/cluster.cgi?seq=CS0DB006B10QPl&cluster=4390.r. Contact :  
Feng Liang Email : fliang@lifetech.com URL :  
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600

BX352919 1201 bp mRNA linear EST 02-MAY-2003  
Homo sapiens NEUROBLASTOMA COT 10-NORMALIZED Homo sapiens  
CDNA clone CS0DB006YC20 5-PRIME, mRNA sequence.  
BX352919 Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 1201)  
Li,W.B., Gruber,C., Jessee,J. and Polayes,D.  
Full-length cDNA libraries and normalization  
Unpublished  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of  
Invitrogen. This sequence belongs to sequence cluster 4390.r For  
more information about this cluster, see  
http://www.genoscope.cns.fr/  
cgi-bin/cluster.cgi?seq=CS0DB006B10QPl&cluster=4390.r. Contact :  
Feng Liang Email : fliang@lifetech.com URL :  
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600

BM923064 AGENCOURT  
AL525111 AL525111  
BI837148 603089920  
BI827333 603077811  
BM014974 603640923  
AU141625 AU141625  
BX370424 BX370424  
BU181305 AGENCOURT  
BM473687 AGENCOURT  
BQ436653 AGENCOURT  
BI253559 602973485  
AL522085 AL522085  
BM010281 603631154  
BI818555 603033078  
AI963473 WT6302.X  
BG393676 602412094  
BF797243 602577745  
BM45513 AGENCOURT  
BU172522 AGENCOURT  
BM465183 AGENCOURT  
BG282890 602406016  
AK002912 Mus muscu  
AL556719 AL556719  
BM019985 603648507  
AU280528 AU280528  
BE621964 601440609  
CB182842 AGENCOURT  
BM424026 AGENCOURT  
BY703409 BY703409  
AK013214 Mus muscu  
BQ571032 UI-M-FB0-  
BE267359 601189568  
BQ964436 AGENCOURT  
BI254417 602976749  
BI415573 602988350  
AK004291 Mus muscu  
BM019007 603646981  
BE563597 601334712  
BI329060 602980475  
AL706316 DKFP686C  
BM461077 AGENCOURT

Faraday Avenue Genoscope sequence ID : CS0DB006BB10QP1.

## FEATURES

Location/Qualifiers  
1..1201  
/organism="Homo sapiens"  
/mol\_type="mrna"  
/db\_xref="taxon:9606"  
/clone="CS0DB006Y20"  
/tissue\_type="NEUROBLASTOMA COT 10-NORMALIZED"  
/clone\_lib="Homo sapiens NEUROBLASTOMA COT 10-NORMALIZED"  
/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."  
BASE COUNT 214 a 355 c 367 g 226 t 39 others  
ORIGIN  
Query Match 73.3%; Score 980.2; DB 13; Length 1201;  
Best Local Similarity 96.2%; Pred. No. 2.2e-188;  
Matches 1024; Conservative 19; Mismatches 17; Indels 5; Gaps 4;  
QY 18 CGAGCAACAGATGGCAGCTGGAGCGCTCTCCGGCGCGCTTCGTGGCGCG 77  
DB 68 CGAGCAACAGATGGCAGCTGGCGGA--CCGTCTCCGGCGCGCTTCGTGGCGCG 126  
QY 78 GCCGCATGGGGGGCCATCGCGAGGGCTCATCAGACGAGGAAAAGTGAAGCTCAG 137  
DB 127 GCGCGCATGGCGGGGCCATCGCGCAGGGCTCATCAGACGAGAAAAGTGAAGCTCAG 186  
QY 138 CACATACGGCCAGTGACCAACAGACAGAACTATGTACATTTCAAGCTCTGGTTGC 197  
DB 187 CACATACGGCCAGTGACCAACAGACAGAACTATGTACATTTCAAGCTCTGGTTGC 246  
QY 198 CGGACACGCACTCCACACGAGAGGTGTCGAGAGTGCCTGCTCGTCATCTTTGCGACC 257  
DB 247 CGGACACGCACTCCACACGAGAGGTGTCGAGAGTGCCTGCTCGTCATCTTTGCGACC 306  
QY 258 AAGCCTATGTGTCGCGAGTGTCTCGCAGAGTGGCTCCTGTGTACCACTGAACAC 317  
DB 307 AAGCCTATGTGTCGCGAGTGTCTCGCAGAGTGGCTCCTGTGTACCACTGAACAC 366  
QY 318 ATCTGGTGTCCGTGGTGTCTGAGACCTCGTGGAGAGTGTCTGCGCCCA 377  
DB 367 ATCTGGTGTCCGTGGTGTCTGAGACCTCGTGGAGAGTGTCTGCGCCCA 426  
QY 378 AACACAGGGTGTGCGGGTCTTCCCAAGCTGTGTGTGTCAGGAAGGGCCATA 437  
DB 427 AACACAGGGTGTGCGGGTCTTCCCAAGCTGTGTGTGTCAGGAAGGGCCATA 486  
QY 438 GTGATGGCGCGGGCGCCACGTGGGAGCAGGAGACCAAGCTCTGTGAGCATCTGCTG 497  
DB 487 GTGATGGCGCGGGCGCCACGTGGGAGCAGGAGACCAAGCTCTGTGAGCATCTGCTG 546  
QY 498 GAGGCTGTGGGGGTGTGAGGAGTGTCTGAGACCTACGTGACATTCACACTGGCCCTC 557  
DB 547 GAGGCTGTGGGGGTGTGAGGAGTGTCTGAGACCTACGTGACATTCACACTGGCCCTC 606  
QY 558 AGTGACGTGGCTGGCTGTGTGTGTCATCTTCGAGGCCCTGTGTAAGGAGCGTC 617  
DB 607 AGTGACGTGGCTGGCTGTGTGTGTCATCTTCGAGGCCCTGTGTAAGGAGCGTC 666  
QY 618 AGATGGGCATGCCAG--CAGCCTGGCCCCACCGCATGCTGTGCCAGACCTGTGTGGGAC 676  
DB 667 AAGATGGGCATGCCAGACGCTGGCCCCACCGCATGCTGTGCCAGACCTGTGTGGGAC 726  
QY 677 GGCACAGATGTCTGCAGGAGGCGCAACACCGCCAGCTCGCTCAGACGTGTGAC 736  
DB 727 GGCACAGATGTCTGCAGGAGGCGCAACACCGCCAGCTCGCTCAGACGTGTGAC 786  
QY 737 CCGGGTGGCACCACTATCTATGACTTCCACGCTTGGAGAGGGGGGTGTGAGCAGC 796  
DB 787 CCGGGTGGCACCACTATCTATGACTTCCACGCTTGGAGAGGGGGGTGTGAGCAGC 846  
QY 797 CACCATGAGCGCGTGGAGGCTGCCACCTGCCGGGCCAAGAGCTCAGCAGAAAGTAGGC 856

DB 847 CACCATGAGCGCGTGGAGGCTGCCACCTGCCGGGCAAGAGCTCAGCAGAAAGTAGGC 906  
QY 857 TGGGCTCTGGCATCCCTTTCTCTGCTGTGCCCCCTGCTCTCCCTGCTGCTTCCCT 916  
DB 907 TGGGCTCTGGCATCCCTTTCTCTGCTGTGCCCCCTGCTCTCCCTGCTGCTTCCCT 966  
QY 917 GAGGACTGCGGCTCCCTCTCTGCTGCTGATGAGGTCTCTCTCTCTCTCTCTCTCT 976  
DB 967 GAGGACTGCGGCTCCCTCTCTGCTGCTGATGAGGTCTCTCTCTCTCTCTCTCTCT 1026  
QY 977 CAGGAATGCGGCGGCGAGGCTGGAGGTTCAGAGGCGGGGAGCCCGACCACT 1036  
DB 1027 MAGGGAATGAGGGGAGGCTTGGAGGTTCAGAGGCTTCAGAGGCGGGGAGCCGACATG 1085  
QY 1037 GGGGACACT 1081  
DB 1086 GGGR--AATCYCCTCCCAATKRAAAAAGGACGTGGGKGGT 1128  
AL519338 1201 bp mRNA linear EST 22-MAY-2003  
AL519338 Homo sapiens NEUROBLASTOMA COT 10-NORMALIZED Homo sapiens  
CDNA clone CS0DB004YF19 5-PRIME, mRNA sequence.  
AL519338  
AL519338.2 GI:31037685  
EST.  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
1 (bases 1 to 1201)  
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
Full-length cDNA libraries and normalization  
Unpublished  
On Feb 13, 2001 this sequence version replaced gi:12782831.  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 Evry cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of  
Invitrogen. This sequence belongs to sequence cluster 4390.r For  
more information about this cluster, see  
http://www.genoscope.cns.fr/  
cgi-bin/cluster.cgi?seq=CS0DB004YF19&cluster=4390.F. Contact :  
Feng Liang Email : fliang@lifetech.com URL :  
http://fulllength.invitrogen.com/InvitroGen Corporation 1600  
Faraday Avenue Genoscope sequence ID : CS0DB004YF19QPI.  
FEATURES  
Source  
1..1201  
/organism="Homo sapiens"  
/mol\_type="mrna"  
/db\_xref="taxon:9606"  
/clone="CS0DB004YF19"  
/tissue\_type="NEUROBLASTOMA COT 10-NORMALIZED"  
/clone\_lib="Homo sapiens NEUROBLASTOMA COT 10-NORMALIZED"  
/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."  
BASE COUNT 215 a 346 c 372 g 224 t 44 others  
ORIGIN  
Query Match 66.5%; Score 889.2; DB 9; Length 1201;  
Best Local Similarity 98.1%; Pred. No. 5.9e-170;  
Matches 952; Conservative 9; Mismatches 3; Indels 6; Gaps 6;  
QY 23 CAACAAGATGGCAGTGGAGCGCTCTCGCGCGCGCTGGGCTGTGTGGGCGCGCGCG 82  
DB 67 CAACAAGATGGCAGTGGAG--CCGTCTCGCGCGCGCTGGGCTGTGTGGGCGCGCGCG 125  
QY 83 CATGGCGGGGCCATCGCGAGGCTCTATCAGCAGGAGGAAAAGTGAAGCTCAGACAT 142



Db 445 GATGGCGGGGGCCGACAGTGGGAGACGAGAGACCAACCTCTCTGACGATCTGCTGGA 504  
QY 500 GGCCTGTGGGCGGTGTCAGGAGGTGCTGAAGCTTACGTGCGACATCCACACTGGCTCAG 559  
Db 505 GGCCTGTGGGCGGTGTCAGGAGGTGCTGAAGCTTACGTGCGACATCCACACTGGCTCAG 564  
QY 560 TGGCAGTGGGCGGTGTCAGTGTGTCATCTCCGAGGCGCTGGCTGAAGAGGCGGTCAA 619  
Db 565 TGGCAGTGGGCGGTGTCAGTGTGTCATCTCCGAGGCGCTGGCTGAAGAGGCGGTCAA 624  
QY 620 GATGGGATGCCAGAGCGCTGGCCACCGCATCGCTGCCAGACCGCTGCTGGGACGGC 679  
Db 625 GATGGGATGCCAGAGCGCTGGCCACCGCATCGCTGCCAGACCGCTGCTGGGACGGC 684  
QY 680 CAAGATGCTGTGTCAGCAGGGCCAAACACCCAGGCCAGCTGCGCTCAGAGTGTGCACCC 739  
Db 685 CAAGATGCTGTGTCAGCAGGGCCAAACACCCAGGCCAGCTGCGCTCAGAGTGTGCACCC 744  
QY 740 GGGTGGCACCAACCATCTATGACTTCCACGCGCTGGAGCAGGGCGGCTGCGAGCAGCCAC 799  
Db 745 GGGTGGCACCAACCATCTATGACTTCCACGCGCTGGAGCAGGGCGGCTGCGAGCAGCCAC 804  
QY 800 CATGAGCGCGCTGAGAGTGCACACTCCCGGGCCAGAGAGCTCAGCAGAAAGTA-GGCTG 858  
Db 805 CATGAGCGCGCTGAGAGTGCACACTCCCGGGCCAGAGAGCTCAGCAGAAAGTAGGGCTG 864  
QY 859 GGCTGTGGCCAT-CCTTTCTGCTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 916  
Db 865 GGCTGTGGCCAT-CCTTTCTGCTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 924  
QY 917 GAGGACTGC-GGCTCCCTCCCTCTCTGATGAGG-TCCTCTACTGCTCTCTCTCTCTCTCT 973  
Db 925 GAGGACTCGGGTCCCTCTCTCTGATGAGGAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 984  
QY 974 GCACAGGAAA 984  
Db 985 GCACAGGAAA 995

RESULT 4  
AL526237  
LOCUS  
DEFINITION AL526237 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED EST 23-MAY-2003  
ACCESSION AL526237  
VERSION AL526237  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
On Feb 13, 2001 this sequence version replaced gi:12789730.  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
Bp 191 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of  
Invitrogen. This sequence belongs to sequence cluster 4390.r For  
more information about this cluster, see  
http://www.genoscope.cns.fr/  
cgi-bin/Cluster.cgi?seq=CS0DC016V13&cluster=4390.r. Contact :  
Feng Liang Email: fliang@lifetech.com URL :  
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600  
Faraday Avenue Genoscope sequence ID : CS0DC016AD07Qp1.  
Location/Qualifiers  
1. .977  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"

/clone="CS0DC016V13"  
/tissue\_type="NEUROBLASTOMA COT 25-NORMALIZED"  
/clone\_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"  
/note="1st strand cDNA was primed with a NotI-oligo(dT)  
primer. Five prime end enriched, double-strand cDNA was  
digested with Not I and cloned into the Not I and EcoR V  
sites of the PCWSPORT 6 vector. Library was normalized."  
BASE COUNT 175 a 312 c 316 g 173 t 1 others  
ORIGIN  
Query Match  
Best Local Similarity 65.8%; Score 880.6; DB 9; Length 977;  
Matches 916; Conservative 1; Mismatches 5; Indels 3; Gaps 3;  
QY 7 CGCAGTCTGTCCAGGACCAACAGATGGCAGCTGCGAGCGGTCTCCGGCGCGCTGGGCT 66  
Db 56 CCGGGGATGTCCAGGAGCAACAGATGGCAGCTGCGGA-CCGTCTCCGGCGCGCTGGGCT 114  
QY 67 TCGTGGCGGGGGCGCATGCGGGGGCCATCGCGCAGGGCCCTCATCAGAGCAGGAAAAG 126  
Db 115 TCGTGGCGGGGGCGCATGCGGGGGCCATCGCGCAGGGCCCTCATCAGAGCAGGAAAAG 174  
QY 127 TGGAGCTCAGCATACTACTGCGCAGTGCACCAACACAGAGAACCTATGCTACTTTCAAG 186  
Db 175 TGGAGCTCAGCATACTACTGCGCAGTGCACCAACACAGAGAACCTATGCTACTTTCAAG 234  
QY 187 CTCTGGCTTGGCGGACCAACACTCCAAACAGAGAGTGTGAGAGCTGCTGCTCGTCA 246  
Db 235 CTCTGGCTTGGCGGACCAACACTCCAAACAGAGAGTGTGAGAGCTGCTGCTCGTCA 294  
QY 247 TCTTTGGCACACAGCCTCATGTGCTGCGAGCTCTCTGCGAGAGTGGCTCTCTGTGCTCA 306  
Db 295 TCTTTGGCACACAGCCTCATGTGCTGCGAGCTCTCTGCGAGAGTGGCTCTCTGTGCTCA 354  
QY 307 CCAGTGAACACATCTTGGTCTCGCTGGGTGCTCTGAGACACCTCGAGAGGAGC 366  
Db 355 CCAGTGAACACATCTTGGTCTCGCTGGGTGCTCTGAGACACCTCGAGAGGAGC 414  
QY 367 TGCTGCCCCCAACACACAGGCTGCTGGGGTCTTTGCCCAACCTGCTGCTGTGCTCCAGG 426  
Db 415 TGCTGCCCCCAACACACAGGCTGCTGGGGTCTTTGCCCAACCTGCTGCTGTGCTCCAGG 474  
QY 427 AAGGGGCATAGTATGATGGCGGGGGCGCCACAGTGGGAGCAGCAGAGCAAGCTCTGTC 486  
Db 475 AAGGGGCATAGTATGATGGCGGGGGCGCCACAGTGGGAGCAGCAGAGCAAGCTCTGTC 534  
QY 487 AGCATCTGTGGAGGCTGTGGGGTGTGAGAGAGTGGCTGAAGCTTACGCTGCAATCC 546  
Db 535 AGCATCTGTGGAGGCTGTGGGGTGTGAGAGAGTGGCTGAAGCTTACGCTGCAATCC 594  
QY 547 ACAGTGGCCTCAGTGGCAGTGGGCTTCCGTGCTGTGCTATCTCCGAGGCGCTGGCTG 606  
Db 595 ACAGTGGCCTCAGTGGCAGTGGGCTTCCGTGCTGTGCTATCTCCGAGGCGCTGGCTG 654  
QY 607 AAGGAGCGCTCAGATGGGATGCCAGAGCTGGCCACCGCATCGCTGCGCAGACCC 666  
Db 655 AAGGAGCGCTCAGATGGGATGCCAGAGCTGGCCACCGCATCGCTGCGCAGACCC 714  
QY 667 TGCTGGGAGCGGCCAAGATGCTGTCACAGGGGCCAACACCCAGCCAGCTGCTGCTCAG 726  
Db 715 TGCTGGGAGCGGCCAAGATGCTGTCACAGGGGCCAACACCCAGCCAGCTGCTGCTCAG 774  
QY 727 ACAGTGTGACCCCGGGTGGCACCACCATCATGAGTCCACAGCCCTGGAGAGGCGGGC 786  
Db 775 ACAGTGTGACCCCGGGTGGCACCACCATCATGAGTCCACAGCCCTGGAGAGGCGGGC 834  
QY 787 TGGCAGAGCACCACATGAGCGCTGCCAGCTGCCAGGCGGCCAACAGGAGCTCAGCA 846  
Db 835 TGGCAGAGCACCACATGAGCGCTGCCAGCTGCCAGGCGGCCAACAGGAGCTCAGCA 893  
QY 847 GAAAGTAGGCTGGGCTCTGGCCATTCCTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 906  
Db 894 GAAAGTAGGCTGGGCTCTGGCCATTCCTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 953



Unpublished  
On Feb 13, 2001 this sequence version replaced gi:12788604.  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web: [www.genoscope.cns.fr](http://www.genoscope.cns.fr)  
Library was constructed by Life Technologies, a division of  
Invitrogen. This sequence belongs to sequence cluster 4390.r For  
more information about this cluster, see  
<http://www.genoscope.cns.fr/>  
<http://cgb-bin/cluster.cgi?seq=CS0DC005BF03QPl&cluster=4390.r>. Contact :  
Feng Liang Email : [liang@lifetech.com](mailto:liang@lifetech.com) URL :  
<http://fulllength.invitrogen.com/> Invitrogen Corporation 1600  
Faraday Avenue Genoscope sequence ID : CS0DC005BF03QPl.

| FEATURES                | SOURCE     |
|-------------------------|------------|
| 1. <b>Feature 1</b>     | Source 1   |
| 2. <b>Feature 2</b>     | Source 2   |
| 3. <b>Feature 3</b>     | Source 3   |
| 4. <b>Feature 4</b>     | Source 4   |
| 5. <b>Feature 5</b>     | Source 5   |
| 6. <b>Feature 6</b>     | Source 6   |
| 7. <b>Feature 7</b>     | Source 7   |
| 8. <b>Feature 8</b>     | Source 8   |
| 9. <b>Feature 9</b>     | Source 9   |
| 10. <b>Feature 10</b>   | Source 10  |
| 11. <b>Feature 11</b>   | Source 11  |
| 12. <b>Feature 12</b>   | Source 12  |
| 13. <b>Feature 13</b>   | Source 13  |
| 14. <b>Feature 14</b>   | Source 14  |
| 15. <b>Feature 15</b>   | Source 15  |
| 16. <b>Feature 16</b>   | Source 16  |
| 17. <b>Feature 17</b>   | Source 17  |
| 18. <b>Feature 18</b>   | Source 18  |
| 19. <b>Feature 19</b>   | Source 19  |
| 20. <b>Feature 20</b>   | Source 20  |
| 21. <b>Feature 21</b>   | Source 21  |
| 22. <b>Feature 22</b>   | Source 22  |
| 23. <b>Feature 23</b>   | Source 23  |
| 24. <b>Feature 24</b>   | Source 24  |
| 25. <b>Feature 25</b>   | Source 25  |
| 26. <b>Feature 26</b>   | Source 26  |
| 27. <b>Feature 27</b>   | Source 27  |
| 28. <b>Feature 28</b>   | Source 28  |
| 29. <b>Feature 29</b>   | Source 29  |
| 30. <b>Feature 30</b>   | Source 30  |
| 31. <b>Feature 31</b>   | Source 31  |
| 32. <b>Feature 32</b>   | Source 32  |
| 33. <b>Feature 33</b>   | Source 33  |
| 34. <b>Feature 34</b>   | Source 34  |
| 35. <b>Feature 35</b>   | Source 35  |
| 36. <b>Feature 36</b>   | Source 36  |
| 37. <b>Feature 37</b>   | Source 37  |
| 38. <b>Feature 38</b>   | Source 38  |
| 39. <b>Feature 39</b>   | Source 39  |
| 40. <b>Feature 40</b>   | Source 40  |
| 41. <b>Feature 41</b>   | Source 41  |
| 42. <b>Feature 42</b>   | Source 42  |
| 43. <b>Feature 43</b>   | Source 43  |
| 44. <b>Feature 44</b>   | Source 44  |
| 45. <b>Feature 45</b>   | Source 45  |
| 46. <b>Feature 46</b>   | Source 46  |
| 47. <b>Feature 47</b>   | Source 47  |
| 48. <b>Feature 48</b>   | Source 48  |
| 49. <b>Feature 49</b>   | Source 49  |
| 50. <b>Feature 50</b>   | Source 50  |
| 51. <b>Feature 51</b>   | Source 51  |
| 52. <b>Feature 52</b>   | Source 52  |
| 53. <b>Feature 53</b>   | Source 53  |
| 54. <b>Feature 54</b>   | Source 54  |
| 55. <b>Feature 55</b>   | Source 55  |
| 56. <b>Feature 56</b>   | Source 56  |
| 57. <b>Feature 57</b>   | Source 57  |
| 58. <b>Feature 58</b>   | Source 58  |
| 59. <b>Feature 59</b>   | Source 59  |
| 60. <b>Feature 60</b>   | Source 60  |
| 61. <b>Feature 61</b>   | Source 61  |
| 62. <b>Feature 62</b>   | Source 62  |
| 63. <b>Feature 63</b>   | Source 63  |
| 64. <b>Feature 64</b>   | Source 64  |
| 65. <b>Feature 65</b>   | Source 65  |
| 66. <b>Feature 66</b>   | Source 66  |
| 67. <b>Feature 67</b>   | Source 67  |
| 68. <b>Feature 68</b>   | Source 68  |
| 69. <b>Feature 69</b>   | Source 69  |
| 70. <b>Feature 70</b>   | Source 70  |
| 71. <b>Feature 71</b>   | Source 71  |
| 72. <b>Feature 72</b>   | Source 72  |
| 73. <b>Feature 73</b>   | Source 73  |
| 74. <b>Feature 74</b>   | Source 74  |
| 75. <b>Feature 75</b>   | Source 75  |
| 76. <b>Feature 76</b>   | Source 76  |
| 77. <b>Feature 77</b>   | Source 77  |
| 78. <b>Feature 78</b>   | Source 78  |
| 79. <b>Feature 79</b>   | Source 79  |
| 80. <b>Feature 80</b>   | Source 80  |
| 81. <b>Feature 81</b>   | Source 81  |
| 82. <b>Feature 82</b>   | Source 82  |
| 83. <b>Feature 83</b>   | Source 83  |
| 84. <b>Feature 84</b>   | Source 84  |
| 85. <b>Feature 85</b>   | Source 85  |
| 86. <b>Feature 86</b>   | Source 86  |
| 87. <b>Feature 87</b>   | Source 87  |
| 88. <b>Feature 88</b>   | Source 88  |
| 89. <b>Feature 89</b>   | Source 89  |
| 90. <b>Feature 90</b>   | Source 90  |
| 91. <b>Feature 91</b>   | Source 91  |
| 92. <b>Feature 92</b>   | Source 92  |
| 93. <b>Feature 93</b>   | Source 93  |
| 94. <b>Feature 94</b>   | Source 94  |
| 95. <b>Feature 95</b>   | Source 95  |
| 96. <b>Feature 96</b>   | Source 96  |
| 97. <b>Feature 97</b>   | Source 97  |
| 98. <b>Feature 98</b>   | Source 98  |
| 99. <b>Feature 99</b>   | Source 99  |
| 100. <b>Feature 100</b> | Source 100 |

| BASE COUNT            | 165 a  | 270 c   | 288 g | 149 t         | 1 others          |
|-----------------------|--------|---|-------|---------------|-------------------|
| ORIGIN                |        |   |       |               |                   |
| Query Match           | 59.48; | Score 794.6;  | DB 9; | Length 873;   |                   |
| Best Local Similarity | 99.84; | Pred. No. 7.7e-151;   |       |               |                   |
| Matches               | 805;   | Conservative  | 1;    | Mismatches 0; | Indels 1; Gaps 1; |
| Qy                    | 16     | TCCGAGGCAACAGAGATGGCAGCTCGGAGGCGCGTCTCCGCGGCGCGGTGGGCTTCGTGGCGG | 75    |               |                   |
| Db                    | 68     | TCCGAGGCAACAGAGATGGCAGCTCGCA-SCGTCTCCGCGCGGTGGGCTTCGTGGCGG      | 126   |               |                   |
| Qy                    | 76     | CGGGCCCATGGCGGGGGCCATCGCGAGGGCCCTCATCAGACGAGGAAAAGTGGAAAGCTC    | 135   |               |                   |
| Db                    | 127    | CGGGCCCATGGCGGGGGCCATCGCGAGGGCCCTCATCAGACGAGGAAAAGTGGAAAGCTC    | 186   |               |                   |
| Qy                    | 136    | AGCACATCTGGCGCAGTGCACCAACAGACAGGAACCTTATGTCATTTTCAAGCTCTGGGTT   | 195   |               |                   |
| Db                    | 187    | AGCACATCTGGCGCAGTGCACCAACAGACAGGAACCTTATGTCATTTTCAAGCTCTGGGTT   | 246   |               |                   |
| Qy                    | 196    | GCGGACCCAGCGACTCCCAACAGGAGTGCTGCAGAGCTGCCTGCTCGTTCATCTTTGCCA    | 255   |               |                   |
| Db                    | 247    | GCGGACCCAGCGACTCCCAACAGGAGTGCTGCAGAGCTGCCTGCTCGTTCATCTTTGCCA    | 306   |               |                   |
| Qy                    | 256    | CCAAGCCTCATGTGCTGCCAGCTGTCTTGGCAGAGTGGTCTCTGTGGTCAACCACTGAAC    | 315   |               |                   |
| Db                    | 307    | CCAAGCCTCATGTGCTGCCAGCTGTCTTGGCAGAGTGGTCTCTGTGGTCAACCACTGAAC    | 366   |               |                   |
| Qy                    | 316    | ACATCTTGGTGTCCGTGGCTGCTGGGTGTCTTGAGCACCCCTGGAGAGCTGCTGCCCC      | 375   |               |                   |
| Db                    | 367    | ACATCTTGGTGTCCGTGGCTGCTGGGTGTCTTGAGCACCCCTGGAGAGCTGCTGCCCC      | 426   |               |                   |
| Qy                    | 376    | CAAAACACAGGGTGCTGCGGGTCTTGCCCAACCTGCCCTGTGTGGTCCAGGAAGGGCCCA    | 435   |               |                   |
| Db                    | 427    | CAAAACACAGGGTGCTGCGGGTCTTGCCCAACCTGCCCTGTGTGGTCCAGGAAGGGCCCA    | 486   |               |                   |
| Qy                    | 436    | TAGTGATGGCGGGGGCCGACGCTGGGGAGCAGCAGACCAAGTCTCTGACGATCTGC        | 495   |               |                   |
| Db                    | 487    | TAGTGATGGCGGGGGCCGACGCTGGGGAGCAGCAGACCAAGTCTCTGACGATCTGC        | 546   |               |                   |
| Qy                    | 496    | TGGAGGCCCTGTGGGCGGTGTGAGGAGGTGCCTGAAGCCTACGTCGACATCCACACATGGCC  | 555   |               |                   |
| Db                    | 547    | TGGAGGCCCTGTGGGCGGTGTGAGGAGGTGCCTGAAGCCTACGTCGACATCCACACATGGCC  | 606   |               |                   |
| Qy                    | 556    | TCAGTGGCAGTGGCTGGCCCTTCGTGTGTGCAATTCGAGAGGCCCTGGCTGAAGAGCCG     | 615   |               |                   |
| Db                    | 607    | TCAGTGGCAGTGGCTGGCCCTTCGTGTGTGCAATTCGAGAGGCCCTGGCTGAAGAGCCG     | 666   |               |                   |

## QY

|    |     |   |     |
|----|-----|---|-----|
| Qy | 616 | TCAGATGGGCATGCCAGCAGCTGGCCCAACCGCATCGTGTGCCCAAGACCCCTCTGSGGA  | 675 |
| Db | 667 | TCAGATGGGCATGCCAGCAGCTGGCCCAACCGCATCGTGTGCCCAAGACCCCTCTGSGGA  | 726 |
| Qy | 676 | CGGCCAAGATGCTGTGTCACGAGGGCCAAACACCAGCCCGCAGCTGCCTCAGAGCTGTGCA | 735 |
| Db | 727 | CGGCCAAGATGCTGTGTCACGAGGGCCAAACACCAGCCCGCAGCTGCCTCAGAGCTGTGCA | 786 |
| Qy | 736 | CCCCGGTGGCAACCAACCATCTATGGACATCCACGCCCTGGAGCAGGGCGGGCTCGAGCAG | 795 |
| Db | 787 | CCCCGGTGGCAACCAACCATCTATGGACATCCACGCCCTGGAGCAGGGCGGGCTCGAGCAG | 846 |
| Qy | 796 | CCACCATGAGCCCGTGGAGGCTGCCA                                    | 822 |
| Db | 847 | CCACCATGAGCCCGTGGAGGCTGCCA                                    | 873 |

RESULT 7  
BI837148

|            |  |        |        |                 |
|------------|--|--------|--------|-----------------|
| LOCUS      | BI37148  | 873 bp | linear | EST 04-Oct-2000 |
| DEFINITION | 60308920F1.NIH.MGC_120 Homo sapiens cDNA clone IMAGE:5228956 5', |        |        |                 |
| ACCESSION  | U030892.1  |        |        |                 |
| VERSION    | BI37148  |        |        |                 |
| KEYWORDS   | EST  |        |        |                 |
| SOURCE     | Homo sapiens   |        |        |                 |
| ORGANISM   | Homo sapiens (human)   |        |        |                 |

| REFERENCE | AUTHORS | TITLE | JOURNAL | COMMENT |
|-----------|---------|-------|---------|---------|
| 1         | ...     | ...   | ...     | ...     |
| 2         | ...     | ...   | ...     | ...     |
| 3         | ...     | ...   | ...     | ...     |
| 4         | ...     | ...   | ...     | ...     |
| 5         | ...     | ...   | ...     | ...     |
| 6         | ...     | ...   | ...     | ...     |
| 7         | ...     | ...   | ...     | ...     |
| 8         | ...     | ...   | ...     | ...     |
| 9         | ...     | ...   | ...     | ...     |
| 10        | ...     | ...   | ...     | ...     |
| 11        | ...     | ...   | ...     | ...     |
| 12        | ...     | ...   | ...     | ...     |
| 13        | ...     | ...   | ...     | ...     |
| 14        | ...     | ...   | ...     | ...     |
| 15        | ...     | ...   | ...     | ...     |
| 16        | ...     | ...   | ...     | ...     |
| 17        | ...     | ...   | ...     | ...     |
| 18        | ...     | ...   | ...     | ...     |
| 19        | ...     | ...   | ...     | ...     |
| 20        | ...     | ...   | ...     | ...     |
| 21        | ...     | ...   | ...     | ...     |
| 22        | ...     | ...   | ...     | ...     |
| 23        | ...     | ...   | ...     | ...     |
| 24        | ...     | ...   | ...     | ...     |
| 25        | ...     | ...   | ...     | ...     |
| 26        | ...     | ...   | ...     | ...     |
| 27        | ...     | ...   | ...     | ...     |
| 28        | ...     | ...   | ...     | ...     |
| 29        | ...     | ...   | ...     | ...     |
| 30        | ...     | ...   | ...     | ...     |
| 31        | ...     | ...   | ...     | ...     |
| 32        | ...     | ...   | ...     | ...     |
| 33        | ...     | ...   | ...     | ...     |
| 34        | ...     | ...   | ...     | ...     |
| 35        | ...     | ...   | ...     | ...     |
| 36        | ...     | ...   | ...     | ...     |
| 37        | ...     | ...   | ...     | ...     |
| 38        | ...     | ...   | ...     | ...     |
| 39        | ...     | ...   | ...     | ...     |
| 40        | ...     | ...   | ...     | ...     |
| 41        | ...     | ...   | ...     | ...     |
| 42        | ...     | ...   | ...     | ...     |
| 43        | ...     | ...   | ...     | ...     |
| 44        | ...     | ...   | ...     | ...     |
| 45        | ...     | ...   | ...     | ...     |
| 46        | ...     | ...   | ...     | ...     |
| 47        | ...     | ...   | ...     | ...     |
| 48        | ...     | ...   | ...     | ...     |
| 49        | ...     | ...   | ...     | ...     |
| 50        | ...     | ...   | ...     | ...     |
| 51        | ...     | ...   | ...     | ...     |
| 52        | ...     | ...   | ...     | ...     |
| 53        | ...     | ...   | ...     | ...     |
| 54        | ...     | ...   | ...     | ...     |
| 55        | ...     | ...   | ...     | ...     |
| 56        | ...     | ...   | ...     | ...     |
| 57        | ...     | ...   | ...     | ...     |
| 58        | ...     | ...   | ...     | ...     |
| 59        | ...     | ...   | ...     | ...     |
| 60        | ...     | ...   | ...     | ...     |
| 61        | ...     | ...   | ...     | ...     |
| 62        | ...     | ...   | ...     | ...     |
| 63        | ...     | ...   | ...     | ...     |
| 64        | ...     | ...   | ...     | ...     |
| 65        | ...     | ...   | ...     | ...     |
| 66        | ...     | ...   | ...     | ...     |
| 67        | ...     | ...   | ...     | ...     |
| 68        | ...     | ...   | ...     | ...     |
| 69        | ...     | ...   | ...     | ...     |
| 70        | ...     | ...   | ...     | ...     |
| 71        | ...     | ...   | ...     | ...     |
| 72        | ...     | ...   | ...     | ...     |
| 73        | ...     | ...   | ...     | ...     |
| 74        | ...     | ...   | ...     | ...     |
| 75        | ...     | ...   | ...     | ...     |
| 76        | ...     | ...   | ...     | ...     |
| 77        | ...     | ...   | ...     | ...     |
| 78        | ...     | ...   | ...     | ...     |
| 79        | ...     | ...   | ...     | ...     |
| 80        | ...     | ...   | ...     | ...     |
| 81        | ...     | ...   | ...     | ...     |
| 82        | ...     | ...   | ...     | ...     |
| 83        | ...     | ...   | ...     | ...     |
| 84        | ...     | ...   | ...     | ...     |
| 85        | ...     | ...   | ...     | ...     |
| 86        | ...     | ...   | ...     | ...     |
| 87        | ...     | ...   | ...     | ...     |
| 88        | ...     | ...   | ...     | ...     |
| 89        | ...     | ...   | ...     | ...     |
| 90        | ...     | ...   | ...     | ...     |
| 91        | ...     | ...   | ...     | ...     |
| 92        | ...     | ...   | ...     | ...     |
| 93        | ...     | ...   | ...     | ...     |
| 94        | ...     | ...   | ...     | ...     |
| 95        | ...     | ...   | ...     | ...     |
| 96        | ...     | ...   | ...     | ...     |
| 97        | ...     | ...   | ...     | ...     |
| 98        | ...     | ...   | ...     | ...     |
| 99        | ...     | ...   | ...     | ...     |
| 100       | ...     | ...   | ...     | ...     |

CONTACT: Robert Strausberg, Ph.D.  
Email: c9apbs-r@mail.nih.gov  
Tissue Procurement: Life Technologies, Inc.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLAM11576 row: g column: 05  
High quality sequence stop: 863.

FEATURES  
SOURCE

```

1. 1:673
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clones="IMAGE:5228956"
/lab_host="DH10B"
/clone_lib="NIH_MGC_120"
/notes="Organ: pooled pancreas and spleen; Vector:
pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
source anonymous pool of spleen and pancreas from 28 yo
male. Library is oligo-dT primed and directionally cloned
(EcoRV site is destroyed upon cloning). Average insert
size 1.5 kb, insert size range 1-2.5 kb. Library is
normalized and enriched for full-length clones and was
constructed by C. Gruber (Invitrogen). Research Genetics
tracking code 025. Note: this is a NIH_MGC library."
154 a 278 c 290 g 151 t
BASE COUNT

```

BASE CO  
ORIGIN

Query Match 58.8%; Score 786.4; DB 12; Length 873;  
Best Local Similarity 98.1%; Pred. No. 3.5e-149;

| QY | 17 | CCGAGGCAACAAGATGCGAGTGGAGCGGCTCGGGCGC                      | -GTGGGCTCGTGGGCG | 75 |
|----|----|--|------------------|----|
|    |    |  |                  |    |
| Db | 1  | CCGAGGCAACAAGATGCGAGTGGAGCGGCTCGGGCGGCGTCTCGTGGGCG         | 60               |    |
|    |    |  |                  |    |
| QY | 76 | CGGCGCCGATGGCGGGGGCGCATCGCGCAGGGGCTCATCAGAGCAGGAAAGTGAAGTC | 135              |    |

|    |     |   |     |
|----|-----|---|-----|
| Db | 61  | CGCGCCGATGCGGGGGCCATCGCAGAGGCCCTCATCAGACAGAGAAAGTGGAAAGCTC      | 120 |
| Qy | 136 | AGCACAATACTGGCCAGTGCACCAACACAGAGGAACCTATTGTCACTTTCAAGTCTCTGGGTT | 195 |
| Db | 121 | AGCACAATACTGGCCAGTGCACCAACACAGAGGAACCTATTGTCACTTTCAAGTCTCTGGGTT | 180 |
| Qy | 196 | GCCGGACCACGCACCTCCAACACAGAGGTGCTGCAGAGCTGCCTGCCTCATCTTTGGCA     | 255 |
| Db | 181 | GCCGGACCACGCACCTCCAACACAGAGGTGCTGCAGAGCTGCCTGCCTCATCTTTGGCA     | 240 |
| Qy | 256 | CCAAGCCTCATGTGCTGCCAGCTGTCTGGGAGAGGTGGCTCTGTGGTCAACCACTTGAAC    | 315 |
| Db | 241 | CCAAGCCTCATGTGCTGCCAGCTGTCTGGGAGAGGTGGCTCTGTGGTCAACCACTTGAAC    | 300 |
| Qy | 316 | ACATCTTGGTGTGCTGGGTGCTCTGAGCACCCTGGAGGAGCTGCTGCCCC              | 375 |
| Db | 301 | ACATCTTGGTGTGCTGGGTGCTCTGAGCACCCTGGAGGAGCTGCTGCCCC              | 360 |
| Qy | 376 | CAAAACACAGGGTGTGCGGGTCTTTGCCAAACCTGCCCTGTGTGGTCCAGGAAGGGGCCA    | 435 |
| Db | 361 | CAAAACACAGGGTGTGCGGGTCTTTGCCAAACCTGCCCTGTGTGGTCCAGGAAGGGGCCA    | 420 |
| Qy | 436 | TAGTGATGCGCGGGGCCCGCCACGTGGGGAGCAGCAGAGCAACAGCTCCTGCAGCATCTGC   | 495 |
| Db | 421 | TAGTGATGCGCGGGGCCCGCCACGTGGGGAGCAGCAGAGCAACAGCTCCTGCAGCATCTGC   | 480 |
| Qy | 496 | TGGAGCCCTGTGGGCGGTGTGAGAGGTGCCTGAAGCCTACGTGCGATCCACACTGGCC      | 555 |
| Db | 481 | TGGAGCCCTGTGGGCGGTGTGAGAGGTGCCTGAAGCCTACGTGCGATCCACACTGGCC      | 540 |
| Qy | 556 | TCAGTGGCAGTGGCGTGGCCCTTGCTGTGTGCTATTCCTCGAGGCCCTGGCTGAAGAGCGG   | 615 |
| Db | 541 | TCAGTGGCAGTGGCGTGGCCCTTGCTGTGTGCTATTCCTCGAGGCCCTGGCTGAAGAGCGG   | 600 |
| Qy | 616 | TCAAAGTGGGCATGCCACAGCAGCCTTGCGCCACCCGCATCGCTGCCACAGACCTGCTGGGA  | 675 |
| Db | 601 | TCAAAGTGGGCATGCCACAGCAGCCTTGCGCCACCCGCATCGCTGCCACAGACCTGCTGGGA  | 658 |
| Qy | 676 | CGGCCAAGATGTGCTGCACGAGGGCCCAACACCCAGCCCAAGCTGCGCTCAGAGCTGTGCA   | 735 |
| Db | 659 | CGG - CAAGATGCTGTGCACGAGGGCCCAACACCCAGCCCAAGCTGCGCTCAGAGCTGTGCA | 717 |
| Qy | 736 | CCCCGGGTGGCACCAACCATCTATGACATCCACGCCCT - GGAGCAGAGGCGGGCTGCAGCA | 794 |
| Db | 718 | CCCCGGGTGGCACCAACCATCTATGACATCCACGCCCTGGGAGCAGAGTGGGGCTGGCAGCA  | 777 |
| Qy | 795 | GCACCAATGAGCGCGTGGAGGCTGCCACCTTGCCGGGCCAAGGAGCTCAGCAGAAAAGTAG   | 854 |
| Db | 778 | G - CACCATGAGCGCGTGGAGGCTGCCACTTGCCGGCCCAAGGAGCTCCACCAAGAGTAG   | 836 |
| Qy | 855 | GCTGGGCTTGCCCAACCTTTCTGCTGCCCTCTGTGCC                           | 890 |
| Db | 837 | GCTGGGCTTGCCCAACCTTTCTGCTGCCCTCTGTGCC                           | 872 |

|            |   |
|------------|---|
| RESULT 8   |   |
| BI827333   |   |
| LOCUS      | 773 bp mRNA linear EST 04-OCT-2001                                      |
| DEFINITION | 603077811f1 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5169410 5',       |
|            | mRNA sequence.  |
| ACCESSION  | BI827333  |
| VERSION    | BI827333.1 GI:15938883  |
| KEYWORDS   | EST.  |
| SOURCE     | Homo sapiens (human)  |
| ORGANISM   | Homo sapiens  |
|            | Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;                 |
|            | Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.              |
| REFERENCE  | 1 (bases 1 to 773)  |
| AUTHORS    | NIH-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a> . |
| TITLE      | National Institutes of Health, Mammalian Gene Collection (MGC)          |
| JOURNAL    | Unpublished   |
| COMMENT    | Contact: Robert Strausberg, Ph.D.                                       |

```

Db      547 AGTGGCAGTGGCGCTTCGTGTGTGTCATCTCCGAGGCCCTGCTGAAGAGCGCTC 606
Qy      618 AAGATGGGATGCCCGACGAGCTGGCCACCGCATCGCTGCCAGACCCCTGCTGGGAGC 677
Db      607 AAGATGGGATGCCCGACGAGCTGGCCACCGCATCGCTGCCAGACCCCTGCTGGGAGC 666
Qy      678 GCCAAGATGCTGTCGACGAGGCGCCAAACACCCAGCCAGCTGGGCTCAGACGTGTGCAC 737
Db      667 GCCAAGATGCTGTCGACGAGGCGCCAAACACCCAGCCAGCTGGGCTCAGACGTGTGCAC 726
Qy      738 CCGGGTGGCACCACCATCTATGACTCCACGCGCCTGGAGCGGGCG 784
Db      727 CCGGGTGGCACCACCATCTATGACTCCACGCGCCTGGAGCGGG 773

RESULT 9
BM014974
LOCUS   603640923F1 NIH_MGC_87 Homo sapiens cDNA clone IMAGE:5417260 5',
DEFINITION
mRNA sequence.
ACCESSION
VERSION BM014974
KEYWORDS
SOURCE EST.
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 856)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: DCTD/DTF
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LHAM2064 row: m column: 05
High quality sequence stop: 783.

FEATURES
source
1..856
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5417260"
/tissue_type="mammary adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_87"
Note="Organ: breast; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.383 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

BASE COUNT 158 a 267 c 290 g 141 t
ORIGIN

Query Match 56.3%; Score 753; DB 12; Length 856;
Best Local Similarity 97.9%; Pred. No. 2e-142;
Matches 837; Conservative 0; Mismatches 10; Indels 8; Gaps 7;

Qy 18 CGAGGCAACAGATGGCAGCTGGCGAGCGCTCTCCGGCGCGTGGGCTTCGTGGGCGG 77
Db 2 CGAGGCAACAGATGGCAGCTGGCGAGCGCTCTCCGGCGCGTGGGCTTCGTGGGCGG 61
Qy 78 GCGCCGATGGGGGCCCATCGCGCAGGCGCTCATCAGAGCAGGAAAGTGGAGCTCAG 137
Db 62 GCGCCGATGGCGGGGCCCATCGCGCA-GGCTCATCAGAGCAGGAAAGTGGAGCTCAG 120
Qy 138 CACATAGTGGCAGTGCACCAACAGAGCAAGCACTATGTCACCTTCAGCTCGGTTGC 197

```

```

Db      121 CACATAGTGGCAGTGCACCAACAGAGCAAGCAACCTATGTCACTTTCAAGCTCTGGTTGC 180
Qy      198 CGGACCGGCACTCCACACGAGAGGTGCTGCAGAGCTGCCCTGCTCGTCACTCTTTGCCACC 257
Db      181 CGGACCGGCACTCCACACGAGAGGTGCTGCAGAGCTGCCCTGCTCGTCACTCTTTGCCACC 240
Qy      258 AAGCCTCATGTGCTGCCAGCTGTCTTGGCAGAGTGGCTCTGTGTGGTCAACACTGAACAC 317
Db      241 AAGCCTCATGTGCTGCCAGCTGTCTTGGCAGAGTGGCTCTGTGTGGTCAACACTGAACAC 300
Qy      318 ATCTTGGTTCCTGGCTGCTGTGGGTGTCTCTGAGCACCCCTGGAGAGCTGTGTCGCCCA 377
Db      301 ATCTTGGTTCCTGGCTGCTGTGGGTGTCTCTGAGCACCCCTGGAGAGCTGTGTCGCCCA 360
Qy      378 AACACAGGGTGTGCGGGTCTTCCCAACCTGCTGCTGTGGTCCAGGAGGGGCCATA 437
Db      361 AACACAGGGTGTGCGGGTCTTCCCAACCTGCTGCTGTGGTCCAGGAGGGGCCATA 420
Qy      438 GTGATGGCGGGGGGCCGCCACGCTGGGAGCAGCAGCAAGCTCTTGCAGCATCTGCTG 497
Db      421 GTGATGGCGGGGGGCCGCCACGCTGGGAGCAGCAGCAACCTCTCTGCAGCATCTGCTG 480
Qy      498 GAGCCTGTGGCGGGTGTGAGGAGTGCCTGAAGCCTACGTGCACATCCACACTGGCCTC 557
Db      481 GAGCCTGTGGCGGGTGTGAGGAGTGCCTGAAGCCTACGTGCACATCCACACTGGCCTC 540
Qy      558 AGTGGCAGTGGCGCTTTCGTGTGTGTCATTCTCCGAGGCCCTGGCTGAAGAGCGCGTC 617
Db      541 AGTGGCAGTGGCGCTTTCGTGTGTGTCATTCTCCGAGGCCCTGGCTGAAGAGCGCGTC 600
Qy      618 AAGATGGCATGCCAGCAGCTGGGCCACCGCATGCTGCCAGACCCCTGCTGGGGAGCG 677
Db      601 AAGATGGCATGCCAGCAGCTGGGCCACCGCATGCTGCCAGACCCCTGCTGGGGAGCG 659
Qy      678 GCCAAGATGCTGTGCAGAGGCGCCACACCCAGCCAGCTGGCTGCAAGCTGTGCA- 735
Db      660 GCCAAGATGCTGTGCAGAGGCGCCACACCCAGGCCAGCTGGCTGCAAGCTGTGCA- 719
Qy      736 CCGCGGTGGCACCACCATCTAT-TGACTCCAGCCCTGGAGCAGGGC--GGCTCGGAG 792
Db      720 CCGCGGTGGCACCACCATCTATGAGGACTCCAGCCCTGGAGCAGGGCTTGGAG 779
Qy      793 CAGCCACCATGAGGCGCTGGAGCTGCCACCTGCCGGGCCAGGAGCTCAGCAGAAAGT 852
Db      780 CAGCACCATGAGGCGCGGGAGGCTGCAACCTGCCGGG-CAAGAAGGTCAGCAGAAAGT 838
Qy      853 AGGCTGGCTCTGGC 867
Db      839 AGGCTGGCTCTGGC 853

RESULT 10
AUI141625
LOCUS   AUI141625 THYR01 Homo sapiens cDNA clone THYR01000934 5', mRNA
DEFINITION
sequence.
ACCESSION AUI141625
VERSION AUI141625.1 GI:11003146
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 794)
AUTHORS Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,
Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and
Isogai,T.
TITLE HRI human cDNA project
JOURNAL Unpublished
COMMENT Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan

```



Tel: 81-438-52-3975  
Fax: 81-438-52-3986  
Email: genomics@ri.co.jp  
HRI human cDNA project; 5' - & 3'-end one pass sequencing: Helix  
Research Institute; cDNA library construction: Department of  
Virology, Institute of Medical Science, University of Tokyo, and  
Helix Research Institute.

| FEATURES              | source                       | Location/Qualifiers   |
|-----------------------|------------------------------|---|
|                       | 1. .794                      |   |
|                       | /organism="Homo sapiens"     |   |
|                       | /mol_type="mRNA"             |   |
|                       | /db_xref="taxon:9606"        |   |
|                       | /clone="THYROI00934"         |   |
|                       | /tissue_type="thyroid gland" |   |
|                       | /clone_lib="THYROI"          |   |
|                       | /note=vector: pME18SFL3"     |   |
| BASE COUNT            | 142 a 248 c 264 g 133 t      | 7 others  |
| ORIGIN                |                              |   |
| Query Match           | 55.1%;                       | Score 737.6; DB 9; Length 794;                                    |
| Best Local Similarity | 97.7%;                       | Pred. No. 2.6e-139;   |
| Matches               | 765; Conservative            | 0; Mismatches 16; Indels 2; Gaps 2;                               |
| QY                    | 15                           | GTCCGAGGCACACAGATGGCAGCTGCGGAGCCGCTCCGCGGCGCTGGCTTCGTGGGC 74      |
| Db                    | 12                           | GTNCGAGGCACACAGATGGCAGCTGCGGAGCCGCTNTCCGCGGCGCTGGCTTCGTGGGC 71    |
| QY                    | 75                           | CGCGGCCCATGGCGGGGGCCATCGCAGGGCCCTCATCAGACGAGGAAAAGTGAAGCT 134     |
| Db                    | 72                           | CGCGGCCCATGGCGGGGGCCATCGCAGGGCCCTCATCAGACGAGGAAAAGTGAAGCT 131     |
| QY                    | 135                          | CAGCACATACTGGCCAGTGCACCAACACAGACGAGAACTATGTCACTTTCAAGCTCTGGGT 194 |
| Db                    | 132                          | CAGCACATACTGGCCAGTGCACCAACACAGACGAGAACTATGTCACTTTCAAGCTCTGGGT 191 |
| QY                    | 195                          | TGCCGGACACGCACTCCAACAGGAGGTGCTGCAGAGCTGCCTGCTGCTCATCTTTGCC 254    |
| Db                    | 192                          | TGCCGGACACGCACTCCAACAGGAGGTGCTCAAGAGTGCCTGCTGCTCATCTTTGCC 251     |
| QY                    | 255                          | ACCAAGCCTCATGTGCTGCCAGTGTCTCTGGCAGAGTGGCTCCTGTGTCACCACTGAA 314    |
| Db                    | 252                          | ACCAAGCCTCATGTGCTGCCAGTGTCTCTGGCAGAGTGGCTCCTGTGTCACCACTGAA 311    |
| QY                    | 315                          | CACATCTTGGGTTCGGTGGCTGTGGGGTGTCTCTGAGACCCCTGGAGAGTGTGTGCC 374     |
| Db                    | 312                          | CACATCTTGGGTTCGGTGGCTGTGGGGTGTCTCTGAGCACCCCTGGAGAGTGTGTGCC 371    |
| QY                    | 375                          | CCAAACACACGGGTGCTCGGGGTCTTCCCAACTCGCCTGTGTGTCACGAAAGGGGCC 434     |
| Db                    | 372                          | CCAAACACACGGGTGCTCGGGGTCTTCCCAACTCGCCTGTGTGTCACGAAAGGGGCC 431     |
| QY                    | 435                          | ATAGTGTATGGCGGGGGCCGCCAGTGGGAGCAGCAGACCAAGCTCC -TGCAGCATCT 493    |
| Db                    | 432                          | ATAGTGTATGGCGGGGGCCGCCAGTGGGAGCAGCAGACCAAGCTCC -TGCAGCATCT 491    |
| QY                    | 494                          | GCTGAGAGCCTGTGGGCGGTGTGAGAGGTGCCTGAGCCCTACGTCGACATCCACATGG 553    |
| Db                    | 492                          | GCTGAGAGCCTGTGGGCGGTGTGAGAGGTGCCTGAAAGCCTACGTCGACATCCACATGG 551   |
| QY                    | 554                          | CCTCAGTGGCAGTGGCGCTTTCGTGTGTGTGCATTCTCCGAGGGCCCTGGCTGAAGGAGC 613  |
| Db                    | 552                          | CCTCAGTGGCAGTGGCGCTTTCGTGTGTGTGCATTCTCCGAGGGCCCTGGCTGAAGGAGC 611  |
| QY                    | 614                          | CGTCAAGATGGGCATGCCAGCAGCTTGGCCCAACCGCATCGCTGCCAGACCCCTGCTGG 673   |
| Db                    | 612                          | CGTCAAGATGGGCATGCCAGCAGCTTGGCCCAACCGCATCGCTGCCAGACCCCTGCTGG 671   |
| QY                    | 674                          | GACGCCCAAGATGCTGTGTACAGAGGGCCAAACACCAGCCAGCTGCCTCAGAGCTGTG 733    |
| Db                    | 672                          | GACGCCCAAGATGCTGTGTACAGAGGGCCAAACACCAGCCAGCTGTGCTCAGAGCTGTG 731   |
| QY                    | 734                          | CA -CCCCGGGTGGCACCACCATCTATGAGCTTCCACGCCCTGGAGCAGCGCGGCTGCGAG 792 |

```
Db 239 GGCCTGGCTGAAGAGCCGTCGAAGATGGGCATGGCCAGACGCTGGGCCACCGCATCGC 298
Qy 656 TGCCACAGACCTGCTGGGACGCCCAAGATGCTGTGCACGAGGGCCAAACACCCAGCCCA 715
Db 299 TGCCACAGACCTGCTGGGACGCCCAAGATGCTGTGCACGAGGGCCAAACACCCAGCCCA 358
Qy 716 GCTGCGCTCAGAGCTGTGACCCCGGGTGGCACCACCAATCTATGAGTCTCCAGCCCTGGA 775
Db 359 GCTGCGCTCAGAGCTGTGACCCCGGGTGGCACCACCAATCTATGAGTCTCCAGCCCTGGA 418
Qy 776 GCAGGGGGGCTGCGAGAGCAGCACCATGAGCGCGTGGAGGCTGCCACCTGCCGGGCCAA 835
Db 419 GCAGGGGGGCTGCGAGAGCAGCACCATGAGCGCGTGGAGGCTGCCACCTGCCGGGCCAA 478
Qy 836 GGAGCTCAGCAGAAAGTAGGCTGGGCTCGGCATCTCTTCTGCTGCTGTGCCCTGCC 895
Db 479 GGAGCTCAGCAGAAAGTAGGCTGGGCTCTGGCCATCTTCTGCTGTGCCCTGCC 538
Qy 896 TCTCCCTGTGCTCCCTTCCCTGAGGACTCGGCTCCCTCCCTCCCTGATGAGGCTCCT 955
Db 539 TCTCCCTGTGCTCCCTTCCCTGAGGACTCGGCTCCCTCCCTGATGAGGCTCCT 598
Qy 956 ACTGCTCTTCTCCCTTCCCTGAGGACTCGGCTCCCTCCCTCCCTGATGAGGCTCCT 1015
Db 599 ACTGCTCTTCTCCCTTCCCTGAGGAAATGAGGGGCGCAGGACTTGGAGGTTCAAGA 658
Qy 1016 GCGGGGGAGCCCGCAGTGGGACACTCTCCCTCCCTGAGCAGAGGACACCGT 1075
Db 659 GCGGGGGAGCCCGCAGTGGGACACTCTCCCTCCCTGAGCAGAGGACACCGC 718
Qy 1076 GGTGGTGG---CTCTGCTCCCTGCTGCACTGAGCAGCCACCTTCTGCAACATT--GGTCT 1130
Db 719 GGGGGGTGCTCTTGGCCCTTCTGAGAGGGAGCCACCTTCTGCTGCANCTTTGGTTTG 778
Qy 1131 GAGGGGCCCAAGAGATGCTGCTGCTATTTGCGCGCATGTTGGGAGTTGTTGAGG 1190
Db 779 AGNGGGCCCAAGAGAGGGGTTTGGGCCATTTGGCCCATGTTGGGAGTTGTTGAGG 838
Qy 1191 CCATGACAGAACTTACGTTAAACAGCAGCGGTGGCCCAATCCCTGCTGAGAGTGGAG 1250
Db 839 CCATG--ACAGAACTTACGTTAAACAGCAGCGGTGGCCCAATCCCTGCTGAGAGTGGAG 896
Qy 1251 CTGCGCTTGGCTTTTCAAGTGGGCTGTCGAGCTACAGCAGCGGCGGCTGCTCAT 1307
Db 897 CTTGCTCTTTGCTTTTCAAGGCGTCTCGNGGAGATTACAGACGCGGGGTGCTTAAT 953

RESULT 12
BU181305
LOCUS
DEFINITION BU181305 884 bp mRNA linear EST 04-SEP-2002
5', mRNA sequence.
ACCESSION BU181305
VERSION BU181305.1 GI:22695289
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 884)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: DCTD/DPG/Gazdar
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone Distribution: Agencourt Bioscience Corporation
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13196 row: 1 column: 24
```

```
FEATURES
source
High quality sequence stop: 428.
Location/Qualifiers
1..884
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6009767"
/tissue_type="large cell carcinoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_68"
/note="Organ: lung; Vector: pCMV-Sport6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dr.
Average insert size 1.8 kb. Library constructed by Life
Technologies."
BASE COUNT 160 a 281 c 301 g 141 t 1 others
ORIGIN
Query Match 52.7%; Score 704.8; DB 13; Length 884;
Best Local Similarity 96.6%; Pred. No. 1.2e-132;
Matches 762; Conservative 0; Mismatches 23; Indels 4; Gaps 4;
Qy 15 GTCCGAGGACACAAAGATGGCAGCTGCGGAGCCGCTCTCCGCGGCGCTGGCTTCGTGGGC 74
Db 9 GTCCGAGGACACAAAGATGGCAGCTGCGGAGCCGCTCTCCGCGGCGCTGGCTTCGTGGGC 68
Qy 75 GCGGGCGCATGCGGGGGCCATCGCGCAGGGCCCTCATCAGAGCAGGAAAAGTGGAGCT 134
Db 69 GCGGGCGCATGCGGGGGCCATCGCGCAGGGCCCTCATCAGAGCAGGAAAAGTGGAGCT 128
Qy 135 CAGCACATCTAGTGGCCAGTGCACCAACAGACAGAACCTATGTCACTTTCAAGCTCTGGGT 194
Db 129 CAGCACATCTAGTGGCCAGTGCACCAACAGACAGAACCTATGTCACTTTCAAGCTCTGGGT 188
Qy 195 TGCCGGACACGCACTCAACACAGAGGTGCTGCAGAGCTGCCTGCTGCTCATCTTTGCC 254
Db 189 TGCCGGACACGCACTCAACACAGAGGTGCTGCAGAGCTGCCTGCTGCTCATCTTTGCC 248
Qy 255 ACCAAGCCTCATGTGCTGCCAGTGTCTTGGCAGAGTGGCTCTGTGTGCTCACACTGAA 314
Db 249 ACCAAGCCTCATGTGCTGCCAGTGTCTTGGCAGAGTGGCTCTGTGTGCTCACACTGAA 308
Qy 315 CACATCTTGTGTCGCTGGCTGCTGCTGAGCAGCCTTGGAGAGCTGCTGCC 374
Db 309 CACATCTTGTGTCGCTGGCTGCTGCTGAGCAGCCTTGGAGAGCTGCTGCC 368
Qy 375 CCAACACAGGGTGTGCTGCGGGTCTTGCCCAACCTGCTGTGTGCTCCAGGAGGGGCC 434
Db 369 CCAACACAGGGTGTGCTGCGGGTCTTGCCCAACCTGCTGTGTGCTCCAGGAGGGGCC 428
Qy 435 ATAGTGTGGCGGGGGCCGCGCAGTGGGAGCAGCAGCAAGCTTCTGCGAGCATCTG 494
Db 429 ATAGTGTGGCGGGGGCCGCGCAGTGGGAGGTGCCTGAAGCCTACGTCGACATCTG 488
Qy 495 CTGAGGCTGTGGGGCGGTGTGAGGAGGTGCCTGAAGCCTACGTCGACATCTG 554
Db 489 CTGAGGCTGTGGGGCGGTGTGAGGAGGTGCCTGAAGCCTACGTCGACATCTG 548
Qy 555 CTCAGTGGCAGTGGCGTGGCTTCTGCTGTGCAATTTCCGAGGCGCTGGCTGAAGAGCC 614
Db 549 CTCAGTGGCAGTGGCGTGGCTTCTGCTGTGCAATTTCCGAGGCGCTGGCTGAAGAGCC 608
Qy 615 GTCGAAGATGGGATGCGCCAGCAGCTTGCCCA--CCGCACTGCTGCCAGACCTGCTGGG 673
Db 609 GTCGAAGATGGGATGCGCCAGCAGCTTGCCCA--CCGCACTGCTGCCAGACCTGCTGGG 668
Qy 674 GACGGCCAAAGTGTGCTGTGACAGGGCCCAACACCCAGCCAGCTGGCTCAGACGTGTG 733
Db 669 GACGGCCAAAGTGTGCTGTGACAGGGCCCAAC--CCAGCCAGCTGGCTCAGACGTGTG 727
Qy 734 CACCCC--GGGTGGCACCACCATCTAT--GGACTTCCACGCCCTGGAGAGGGCGGCTGCGA 791
Db 728 CACCCC--GGGTGGCACCACCATCTAT--GGACTTCCACGCCCTGGAGAGCGGCGGCTG 787
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 8, 2003, 14:36:31 ; Search time 103 Seconds  
(without alignments)  
5733.695 Million cell updates/sec

Title: US-09-806-536A-29  
Perfect score: 1338  
Sequence: 1 ggtgagcagctcttcga.....ggggcagcagcatatggggt 1338

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_NA:\*  
1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PTUS\_COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length  | ID | Description          |
|------------|-------|-------------|---------|----|----------------------|
| 1          | 201.8 | 15.1        | 1742    | 3  | US-09-099-676-2      |
| 2          | 201.8 | 15.1        | 1742    | 3  | US-09-365-910-2      |
| C 3        | 91.8  | 6.9         | 879     | 4  | US-09-252-991A-359   |
| C 4        | 91.8  | 6.9         | 1164    | 4  | US-09-252-991A-385   |
| 5          | 91.8  | 6.9         | 1419    | 4  | US-09-252-991A-440   |
| C 6        | 88    | 6.5         | 42988   | 4  | US-08-311-731A-128   |
| 7          | 75.8  | 5.7         | 28171   | 4  | US-08-961-527-22     |
| 8          | 73.4  | 5.5         | 4403765 | 3  | US-09-103-840A-2     |
| 9          | 73.4  | 5.5         | 4411529 | 3  | US-09-103-840A-1     |
| 10         | 66.6  | 5.0         | 1147    | 1  | US-08-665-716-1      |
| 11         | 58.6  | 4.4         | 516     | 4  | US-09-252-991A-4140  |
| C 12       | 51.4  | 3.8         | 1116    | 4  | US-09-252-991A-2320  |
| 13         | 51.4  | 3.8         | 2082    | 4  | US-09-252-991A-2596  |
| 14         | 51.4  | 3.8         | 2388    | 4  | US-09-252-991A-2514  |
| 15         | 48.8  | 3.6         | 633     | 4  | US-09-252-991A-488   |
| C 16       | 47.6  | 3.6         | 1686    | 4  | US-09-252-991A-4056  |
| 17         | 46.4  | 3.5         | 1230    | 4  | US-09-252-991A-16191 |
| C 18       | 46.4  | 3.5         | 1812    | 4  | US-09-252-991A-16295 |
| 19         | 46.4  | 3.5         | 2019    | 4  | US-09-252-991A-15983 |
| 20         | 44.6  | 3.3         | 1746    | 4  | US-09-252-991A-4395  |
| C 21       | 44.6  | 3.3         | 2571    | 4  | US-09-252-991A-4351  |
| 22         | 44.6  | 3.3         | 2997    | 4  | US-09-252-991A-4487  |
| C 23       | 44    | 3.3         | 1645    | 4  | US-09-620-312D-807   |
| 24         | 42.8  | 3.2         | 466     | 3  | US-09-154-083-17     |
| 25         | 42.2  | 3.2         | 3978    | 4  | US-09-266-965-19     |
| 26         | 42.2  | 3.2         | 7218    | 1  | US-08-232-463-14     |
| 27         | 42.2  | 3.2         | 12249   | 4  | US-09-266-965-74     |

|      |      |     |       |   |                      |
|------|------|-----|-------|---|----------------------|
| 28   | 42.2 | 3.2 | 18331 | 4 | US-09-266-965-96     |
| C 29 | 42   | 3.1 | 1155  | 4 | US-09-252-991A-12035 |
| C 30 | 42   | 3.1 | 1185  | 4 | US-09-252-991A-12007 |
| 31   | 42   | 3.1 | 1407  | 4 | US-09-232-991A-11979 |
| C 32 | 41.6 | 3.1 | 1365  | 4 | US-09-252-991A-1996  |
| C 33 | 41.6 | 3.1 | 1389  | 4 | US-09-252-991A-1733  |
| C 34 | 41.2 | 3.1 | 870   | 4 | US-09-252-991A-2385  |
| C 35 | 41.2 | 3.1 | 1326  | 4 | US-09-252-991A-2299  |
| C 36 | 41.2 | 3.1 | 1926  | 4 | US-09-249-585A-2     |
| C 37 | 41.2 | 3.1 | 1926  | 4 | US-09-410-399-3      |
| C 38 | 41.2 | 3.1 | 2580  | 3 | US-09-050-863-2      |
| C 39 | 41.2 | 3.1 | 2580  | 4 | US-09-359-081-2      |
| C 40 | 41.2 | 3.1 | 5452  | 2 | US-09-130-114-1      |
| 41   | 41.2 | 3.1 | 8705  | 4 | US-09-647-344A-14    |
| C 42 | 41.2 | 3.1 | 9600  | 3 | US-08-910-647-1      |
| C 43 | 41.2 | 3.1 | 9600  | 4 | US-09-620-925-1      |
| C 44 | 41.2 | 3.1 | 10596 | 1 | US-07-884-811-15     |
| C 45 | 41.2 | 3.1 | 10596 | 1 | US-07-885-971-15     |

## ALIGNMENTS

## RESULT 1

US-09-099-676-2  
; Sequence 2, Application US/09099676  
; Patent No. 6100075  
; GENERAL INFORMATION:  
; APPLICANT: Hillman, Jennifer L.  
; APPLICANT: Corley, Neil C.  
; TITLE OF INVENTION: DELTA 1-PYRROLINE-5-CARBOXYLATE REDUCTASE  
; TITLE OF INVENTION: HOMOLOG  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: Windows  
; SOFTWARE: FastSeq for Windows Version 2.0b  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/099,676  
; FILING DATE: HEREWITH  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Cerrone, Michael C.  
; REGISTRATION NUMBER: 39,132  
; REFERENCE/DOCKET NUMBER: PF-0532 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 650-855-0555  
; TELEFAX: 650-855-0572  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1742 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: PROSNO01  
; CLONE: 2278458  
; US-09-099-676-2

Query Match 15.1%; Score 201.8; DB 3; Length 1742;

```
Best Local Similarity 55.0%; Pred. No. 1.8e-38;
Matches 448; Conservative 0; Mismatches 352; Indels 15; Gaps 2;

QY 42 GAGCGGCTCTCGGGCGCGGCTGCTGCGCGCGCGCGCATGCGGGGGCCATCGCG 101
   ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 180 GAACCGGGACCATGACGCTGGGCTTCATCGGGCGCGCGCGCTATGCTCTGGCG 239

QY 102 CAGGGCTCTACAGAGCAGGAAAGTGAAGCTCAGACATACCTGCGGAGTGCACCAACA 161
   ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 240 CGGGGCTTCACGGCGCAGGATCCTGTCGGCTCAAGAATATAGCCAGCTCCCGAGA 299

QY 162 GACAGGAACC-----TATGTCACTTCAAGCTCTGGGTTGCCGAGCAGGACTCCAAC 215
   ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 300 ATGAACCTGCCACGGTGTCCGGCTCAGGAAGATGGGTGAACCTGACACGCAAC 359

QY 216 CAGGAGTGTGTCAGAGCTGCTGCTGCTATCTTGGCCACCAAGCTCATGTGTGCCA 275
   ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 360 AAGGACGCTGAAGCAGACGACGCTGCTTCTGCTGTGAAGCCATATCATCCGC 419

QY 276 GCTGCTCTGCGAGGTTGCTCTGTTGTCACCACTGAACACATCTTGTGTCCTGGCT 335
   ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 420 TTCACTCTGATGAGATTGGGCGCGAGCTGCAAGCCAGACACATCTGCTCTCTGCG 479

QY 336 GCTGGGTGTCTCTGACGACCTTGGAGGAGCTGCTG-----CCCCAACACACGG 386
   ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 480 GCTGGGTGTCTGACGACCTTGGAGGAGCTGCTGTTGGAAGAAGCTGATGGCAATTCAGCCAGCCGCCAAA 539

QY 387 GTGCTCGGGTCTTGCCCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 446
   ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 540 GTGATTCGCTGTCATGACCAACACCTGTTGTTAGTCAGAGAGGCGCTACAGTGTACGCC 599

QY 447 CCGGGCGGCGACCTGCGGGAGCAGCAGACCAAGCTCTGCGAGCATCTGCTGAGGCTGT 506
   ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 600 ACGGGACCCATGCCCTGTTGGGATGGGAGTGGGAGCTCTCTGAGGAGCTCATGAGCAGCTG 659

QY 507 GGGCGGTGTGAGGAGGTGCTGAAGCTACGTGCGACATCCACACTGCGCTCAGTGCAGT 566
   ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 660 GGCTTCTGCGCTGAGTGAAGGAGCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 719

QY 567 GGGCTGCGCTTGTGTGCTATCTCCGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 626
   ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 720 GGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 779

QY 627 ATCCGAGCAGCTGCGCCACCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 686
   ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 780 TTGCGACGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 839

QY 687 CTGCTGCGAGGGGCAACACCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 746
   ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 840 CTGCTGAGCTGCGAGCAGCATCCATGCCAGCTTAAGGACAATGTCTGCTCCCTGGGGA 899

QY 747 ACCACCATATGAGTCCACGCTTGGAGAGGCGGGCTGCGAGCAGCCACCATGAGC 806
   ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 900 GCACCATCCAGCCCTGCACTTTCTAGAGAGTGGGGCTTCCGCTCTCTGCTCATCAAT 959

QY 807 GCGGTGAGGCTGCGCCACCTGCGCGCGCAAGGAGCT 841
   ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 960 GCAGTTGAGGCTCTGTATCCGAACACGAGAGCT 994
```

## RESULT 2

```
US-09-565-910-2
; Sequence 2, Application US/09565910
; Patent No. 6268192
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Corley, Neil C.
; APPLICANT: Baughn, Mariah R.
; TITLE OF INVENTION: DELTA 1-PYRROLINE-5-CARBOXYLATE REDUCTASE
; TITLE OF INVENTION: HOMOLOG
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
```

```
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/565,910
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/099,676
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Cerrone, Michael C
; REGISTRATION NUMBER: 39,132
; REFERENCE/DOCKET NUMBER: PF-0532 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-855-0572
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1742 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: PROSNON01
; CLONE: 2278458
; US-09-565-910-2
```

Query Match 15.1%; Score 201.8; DB 3; Length 1742;

Best Local Similarity 55.0%; Pred. No. 1.8e-38;

Matches 448; Conservative 0; Mismatches 352; Indels 15; Gaps 2;

```
QY 42 GAGCGGCTCTCGGGCGCGGCTGCTGCGCGCGCGCGCATGCGGGGGCCATCGCG 101
   ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 180 GAACCGGGACCATGACGCTGGGCTTCATCGGGCGCGCGCGCTATGCTCTGGCG 239

QY 102 CAGGGCTCTACAGAGCAGGAAAGTGAAGCTCAGACATACCTGCGGAGTGCACCAACA 161
   ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 240 CGGGGCTTCACGGCGCAGGATCCTGTCGGCTCACAAGATATAGCCAGCTCCCGAGA 299

QY 162 GACAGGAACC-----TATGTCACTTCAAGCTCTGGGTTGCCGAGCAGGACTCCAAC 215
   ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 300 ATGAACCTGCCACGGTGTCCGGCTCAGGAAGATGGGTGAACCTGACACGCAAC 359

QY 216 CAGGAGTGTGTCAGAGCTGCTGCTGCTATCTTGGCCACCAAGCTCATGTGTGCCA 275
   ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 360 AAGGACGCTGAAGCAGACGACGCTGCTTCTGCTGTGAAGCCATATCATCCGC 419

QY 276 GCTGCTCTGCGAGGTTGCTCTGTTGTCACCACTGAACACATCTTGTGTCCTGGCT 335
   ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 420 TTCACTCTGATGAGATTGGGCGCGAGCTGCAAGCCAGACACATCTGCTCTCTGCG 479

QY 336 GCTGGGTGTCTCTGACGACCTTGGAGGAGCTGCTG-----CCCCAACACACGG 386
   ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 480 GCTGGGTGTCTGACGACCTTGGAGGAGCTGCTGTTGGAAGAAGCTGATGGCAATTCAGCCAGCCGCCAAA 539

QY 387 GTGCTCGGGTCTTGCCCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 446
   ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 540 GTGATTCGCTGTCATGACCAACACCTGTTGTTAGTCAGAGAGGCGCTACAGTGTACGCC 599

QY 447 CCGGGCGGCGACCTGCGGGAGCAGCAGACCAAGCTCTGCGAGCATCTGCTGAGGCTGT 506
   ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 600 ACGGGACCCATGCCCTGTTGGGATGGGAGTGGGAGCTCTCTGAGGAGCTCATGAGCAGCTG 659

QY 507 GGGCGGTGTGAGGAGGTGCTGAAGCTACGTGCGACATCCACACTGCGCTCAGTGCAGT 566
```



Db 319 CGGCGCGGCGTATTTCTTCTGCTGATGAGGCCATGACCGAGCGCGGAGAGCTCG 260  
QY 625 GCATGCCAGAGCCTGGCCCCACCGCATCGTGCACACCCCTGCTGGGACGCGCAAGA 684  
Db 259 GCCTGTCCCGGAGACAGCCTCGCGCTGACCTTGACAGCCGCTCGGCGCGGCGCAGA 200  
QY 685 TGCTGTCTGCAGGAGGGCCACACCCAGCCAGCTGCGCTCAGACGTGTGCACCCCGGGTG 744  
Db 199 TGGCGCTGTCCAGGAGGTGAGCGCGCGCAACTGCGCGCGCGGTGACTTCGCCCAACG 140  
QY 745 GCACCAACCATATATGAGACTCCACGCCCTGAGCAGGCGGCTGCGAGCAGCCACCATGA 804  
Db 139 GCACCAAGGAGGCGCGCATCAATCTTCCAGGCCAACGCTTCGAAGCACTGGTGGAGC 80  
QY 805 GCGCGGTGGAGGTGCGCACTCGCGGCGCCAAAGGAGCTCAGCAGAAAG 851  
Db 79 AGGCCCTGAAGCGCGCGCAGCAGCGCTCCGCGAGCTGGCGCAACAG 33

## RESULT 5

US-09-252-991A-440  
; Sequence 440, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 440  
; LENGTH: 1419  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-440

Query Match 6.9%; Score 91.8; DB 4; Length 1419;  
Best Local Similarity 47.4%; Pred. No. 1.2e-12;  
Matches 307; Conservative 0; Mismatches 337; Indels 3; Gaps 1;

QY 208 ACTCCAAACAGAGGTGCTGACAGAGTGCCTGCTGCTATCTTCCACCAAGCCTCATG 267  
Db 761 AGTCCAAACCGAGGCGCGTGGCGGACGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCT 820  
QY 268 TGCTGCCAGTGTCTTGGCAGAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 327  
Db 821 CCATGAAGCCGTGTGCCAAGCCTTGGCGCGCGCTGAAGCGGAGCACTGATCGTCT 880  
QY 328 CGTGTGCTGTGGGTGTCTCTGACACACCTGGAGAGTGTGCTGCCCAACACACACGG 387  
Db 881 CCATGCCCGCGCATCCCTTGGCAGCTGCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCT 940  
QY 388 TGCTGGGTCTTGGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 447  
Db 941 TGGTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1000  
QY 448 GGGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 507  
Db 1001 CCACCGCGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1060  
QY 508 GCGGTGTGAGAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 564  
Db 1061 GCATGCCCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1120  
QY 565 GTGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 624  
Db 1121 GCGGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1180

QY 625 GCATGCCAGAGCCTGGCCCCACCGCATCGTGCACACCCCTGCTGGGACGCGCAAGA 684  
Db 1181 GCCTGTCCCGGAGACAGCCTCGCGCTGACCTTGACAGCCGCTCGGCGGCGCAGA 1240  
QY 685 TGCTGTCTGCAGGAGGGCCACACCCAGCCAGCTGCGCTCAGACGTGTGCACCCCGGGTG 744  
Db 1241 TGGCGCTGTCCAGCAGGTGCGAGCGCGCGCAACTGCGCGCGCGGTGACTTCGCCCAACG 1300  
QY 745 GCACCAACCATATATGAGACTCCACGCCCTGAGCAGGCGGCTGCGAGCAGCCACCATGA 804  
Db 1301 GCACCAACCGAGCGCGCATCAATCTTCCAGGCCAACGCTTCGAAGCACTGGTGGAGC 1360  
QY 805 GCGCGCTGGAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 851  
Db 1361 AGGCCCTGAAGCGCGCAGCAGCGCTCGCGGAGCTGCGCGAAGCAG 1407

## RESULT 6

US-08-311-731A-128/c  
; Sequence 128, Application US/08311731A  
; Patent No. 6583266  
; GENERAL INFORMATION:  
; APPLICANT: SMITH, DOUGLAS  
; APPLICANT: MAO, JEN-I  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES  
; TITLE OF INVENTION: RELATING TO MYCOBACTERIUM TUBERCULOSIS AND LAPRAE FOR  
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS  
; NUMBER OF SEQUENCES: 411  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: WOLF, GREENFIELD & SACKS, P.C.  
; STREET: 600 ATLANTIC AVENUE  
; CITY: BOSTON  
; STATE: MASSACHUSETTS  
; COUNTRY: USA  
; ZIP: 02210  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/311.731A  
; FILING DATE:  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: GATES, EDWARD R.  
; REGISTRATION NUMBER: 31,616  
; REFERENCE/DOCKET NUMBER: C0044/7125  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617/720-3500  
; TELEFAX: 617/720-2441  
; INFORMATION FOR SEQ ID NO: 128:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 42988 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: circular  
; MOLECULE TYPE: DNA (genomic)  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; ORGANISM: MYCOBACTERIUM LEPRAE  
US-08-311-731A-128

Query Match 6.6%; Score 88; DB 4; Length 42988;  
Best Local Similarity 50.7%; Pred. No. 2.3e-11;  
Matches 211; Conservative 0; Mismatches 205; Indels 0; Gaps 0;  
QY 279 GTCTGGCAGAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 338  
Db 27603 GTTCAGGACCGCGCAGTGCAGCAATGACAGCGCTGACAGCTGTTAGTCACCGTGGCTGCT 27544







Db 831 ---AGCATCCGGCGTTCTGAAAGATAAAAGTACTAGCCGGGTGGTGCACGATAGGTG 887  
QY 760 GACTCCACGCCCTGGAGCAGGGCGGCTGCAGACGCCACCATGAGCGCGTGGAGGCTG 819  
Db 888 GGCTGATGGTCTGGAAGAAGAGGGGTGAGAGGGACCGTGGCTAGAGCAGTTAGGGAAG 947  
QY 820 CCACCTCCGGGCGCAAGGAGCTCAGCAGA 848  
Db 948 CTACTGGGTGGCAGTCACCTTGGTAAA 976

RESULT 11  
US-09-252-991A-4140  
; Sequence 4140, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 4140  
; LENGTH: 516  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-4140

Query Match 4.4%; Score 58.6; DB 4; Length 516;  
Best Local Similarity 46.9%; Pred. No. 5.7e-05;  
Matches 220; Conservative 0; Mismatches 244; Indels 5; Gaps 1;

QY 388 TGCTGCGGGTCTTGCCCAACCTGCGCTGTGTGTGTCAGGAGGGGCGCATAGTCATGCGC 447  
Db 36 TGGCGCTGGGGGGCGCGCTCCAGCTGTGGCCACCGCGCGGGGACCTTCGTAG 95  
QY 448 GGGGCCGCCACGTGGGGAGCAGCAGACCAAGCTCTCTGCAGCATCTGCTGGAGGCGCTGTG 507  
Db 96 GTCGGCGCCATGGAGCAGCGCGCGAGGGTCAGGGCGATCAGCAGCGCGGATACGCCGAAT 155  
QY 508 GCGCGGTGTAGGAGGTGCTTGAAGCCTACGTGCACATCCACATGCGCTCAGTGGCAGTG 567  
Db 156 GCAGGTTTGGCGATCTCTGTTCACTCCCGCGCCGAAGCGGCTGTTGATGGTGGCGGCT 215  
QY 568 GCGTGGCTTCGTGTGCAATTCCTCC----GAGGCGCTGGCTGAAGGAGCGCGTCAAGAT 622  
Db 216 TGCTGGCAGCAGCATTTCTTCTGTGTATGAGGCCATACCCAGCGCGCGGAGAGCT 275  
QY 623 GGGCATGCCAGCAGCGCTGGCCACCCATCGTTCGCCAGACCCCTGCTGGGGACGGGCCAA 682  
Db 276 CGGCTGTCCCGGAGACAGCCTCGCGCCTGACCTTCAGACACCCCTCGCGCGGGCGCA 335  
QY 583 GATGCTGCTGCAGAGGGCCCAACACCCAGCCCGAGCTGCGCTCAGAGTGTGACCCGGG 742  
Db 336 GATGGCGCTGTCCAGCAGGTCGAGCGCGGCCGAACCTGCGCGCGGGTGACTTTCGCCCAA 395  
QY 743 TGGCACCACCATCTATGAGTCCACGCGCTTGGAGCAGGGCGGCTCGAGACGCCACCAT 802  
Db 396 CGGCACCCAGCGGCGGATCAATCTTCCAGGCCAACGGTTCCAGACCATGGTGA 455  
QY 803 GAGCGGTGTGAGGCTGCCACCTGCGCGGCCCAAGAGGCTCAGCAGAAAG 851  
Db 456 GCAGGCCCTGAACGCCGCCAGCGCGCTCCGCCGAGCTGGCGGAACAG 504

RESULT 12  
US-09-252-991A-2320/c

; Sequence 2320, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 2320  
; LENGTH: 1116  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-2320

Query Match 3.8%; Score 51.4; DB 4; Length 1116;  
Best Local Similarity 46.1%; Pred. No. 0.0035;  
Matches 172; Conservative 0; Mismatches 201; Indels 0; Gaps 0;

QY 186 GCTCTGGGTTCGGGACCCAGCAGCTCCAAACCCAGGAGGTGCTGCAGAGCTGCTGCTGTC 245  
Db 822 GCCCTGGCCTGCTGTTGCTGTCGTGGCGCGGGAACCCCTGCGCATCTGCGTCCGCGCC 763  
QY 246 ATCTTTGCCACCAAGCCTCATGTGTCGCCAGCTGTCTGCGGAGAGGTGCTCCTGTGGTC 305  
Db 762 CTCGCGGAAGTGCAGCATCGATTCGGGTGGCATCGCTGCACCCCTGCTGGGAGTTC 703  
QY 306 ACCACTGAACACATCTTGGTGTCCGTGGCTGCTGGGTGCTCTCGAGCACCTCGAGGAG 365  
Db 702 ATCATCGGGGTGATTTCGCTGACCGCATCGCCACACCTTCGCCAGTACCTGCTCGAC 643  
QY 366 CTGCTGCCCCCAACACACACGCGGTGCTGCGGGTCTTTGCCCAACCTGCCCTGTGTGCCAG 425  
Db 642 CTGGGACGGGAGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 583  
QY 426 GAAGGGCCCATAGTATGTCGGCGGGGGCGCCACAGTGGGAGAGCAGGACCAAGCTCCTG 485  
Db 582 GGAATGGGATCCCGACCATCCCAACTACATCATCACAGCTCCATCGCCGCGCGCG 523  
QY 486 CAGCATCTCTCGAGGCGCTGTGGGCGGTGTGAGGAGGTGCTGAAAGCCTACGCTCGACATC 545  
Db 522 CTGCTGGAACTGGGGTGCGGCTGATCTCGACATGTTCTTCTACTTCGGCATC 463  
QY 546 CACACTGGCCTCA 558  
Db 462 CTCGCGGACCTCA 450

RESULT 13  
US-09-252-991A-2596  
; Sequence 2596, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 2596  
; LENGTH: 2082  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa





|                           |  |   |       |       |
|---------------------------|--|---|-------|-------|
| BASE COUNT                | 163 a  | 311 c   | 306 g | 152 t |
| ORIGIN                    |  |   |       |       |
| Query Match               | 52.1%; Score 697.4; DB 13; Length 932;   |   |       |       |
| Best Local Similarity     | 97.0%; Pred. No. 3.9e-131;   |   |       |       |
| Matches 732; Conservative | 0; Mismatches 21; Indels 2; Gaps 2;  |   |       |       |
| Qy                        | 41   | GGAGCCGCTCCCGCGCGTGGGCTTCGTGGCGCGGCCGCGCATGGCGGGGCCATCGC      | 100   |       |
| Db                        | 1  | GGAGCCGCTCCCGCGCGTGGGCTTCGTGGCGCGGCCGCGCATGGCGGGGCCATCGC      | 60    |       |
| Qy                        | 101  | GCAGGGCCCTCATCAGACGAGGAAAAGTGAAGCTCAGACATACCTAGCCAGTGCCACCAAC | 160   |       |
| Db                        | 61   | GCAGGGCCCTCATCAGACGAGGAAAAGTGAAGCTCAGACATACCTAGCCAGTGCCACCAAC | 120   |       |
| Qy                        | 161  | AGACAGGAACCTATGTCACTTTCAAGCTCTGGTTCGCGGACACAGCACTCCAAACGAGGA  | 220   |       |
| Db                        | 121  | AGACAGGAACCTATGTCACTTTCAAGCTCTGGTTCGCGGACACAGCACTCCAAACGAGGA  | 180   |       |
| Qy                        | 221  | GGTGTGCAGAGCTGCTGCTGCTATCTTTTGCACCAAGCCTCATGTGCTGCCAGCTGT     | 280   |       |
| Db                        | 181  | GGTGTGCAGAGCTGCTGCTGCTATCTTTTGCACCAAGCCTCATGTGCTGCCAGCTGT     | 240   |       |
| Qy                        | 281  | CCTGGCAGAGTGGCTCCTGTGTGCTCACCCTGAAACATCTTGGTCCGCTGGCTGCTGG    | 340   |       |
| Db                        | 241  | CCTGGCAGAGTGGCTCCTGTGTGCTCACCCTGAAACATCTTGGTCCGCTGGCTGCTGG    | 300   |       |
| Qy                        | 341  | GGTGTCTCTGAGCACCCCTGGAGAGCTGCTGCCGCCCAACACACAGCGGTGCTCGGGTCTT | 400   |       |
| Db                        | 301  | GATGTCTCTGAGCACCCCTGGAGAGCTGCTGCCGCCCAACACACAGCGGTGCTCGGGTCTT | 360   |       |
| Qy                        | 401  | GCCCAACCTGCTGTGTGTGCAGAAAGGCGCATAGTATGATGGCGGGCGCCACAGT       | 460   |       |
| Db                        | 361  | GCCCAACCTGCTGTGTGTGCAGAAAGGCGCATAGTATGATGGCGGGCGCCACAGT       | 420   |       |
| Qy                        | 461  | GGGAGCAGCAGACCAAGCTCTGCAGCATCTGCTGGAGGCTGCTGGCGGTGTGAGGA      | 520   |       |
| Db                        | 421  | GGGAGCAGCAGACCAAGCTCTGCAGCATCTGCTGGAGGCTGCTGGCGGTGTGAGGA      | 480   |       |
| Qy                        | 521  | GGTGCTTGAAGCCTACGTGCATCCACACTGGCCCTCAGTGGCAGTGGCGTGGCCTTGT    | 580   |       |
| Db                        | 481  | GGTGCTTGAAGCCTACGTGCATCCACACTGGCCCTCAGTGGCAGTGGCGTGGCCTTGT    | 540   |       |
| Qy                        | 581  | GTGTGATCTCCGAGGCCCTGGCTGAAGAGCGCTCAAGATGGCGATGCCACAGCCT       | 640   |       |
| Db                        | 541  | GTGTGATCTCCGAGGCCCTGGCTGAAGAGCGCTCAAGATGGCGATGCCACAGCCT       | 600   |       |
| Qy                        | 641  | GGCCACCGCATCGCTGCCAGACCTGCT-GGGGACGCCAAGATGCTGTGCACGAGG       | 699   |       |
| Db                        | 601  | GGCCACCGCATCGCTGCCAGACCTGCTGGGGACCGCAAGATGCTGTGCACGAGG        | 660   |       |
| Qy                        | 700  | GCCAAACCCAGCCAGCTCGCTCAGACGTGTGCACCC-CGGTGGCACCCACCATCTAT     | 758   |       |
| Db                        | 661  | GCCAAACCCAGCCAGCTCGCTCAGACGTGTGCACCCCGGGGGGACCATCATAT         | 720   |       |
| Qy                        | 759  | GGACTCCAGCCCTGGAGCAGGGCGGCTGCCAGC                             | 793   |       |
| Db                        | 721  | GGACTCCCGCCCTGGGAACCAAGGGGGGGGCTGC                            | 755   |       |
| RESULT 15                 |  |   |       |       |
| BI253559                  |  |   |       |       |
| LOCUS                     | 602973485F1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:5112692 5',   |   |       |       |
| DEFINITION                | mRNA sequence.   |   |       |       |
| ACCESSION                 | BI253559   |   |       |       |
| VERSION                   | BI253559.1 GI:14805103   |   |       |       |
| KEYWORDS                  | EST.   |   |       |       |
| SOURCE                    | Homo sapiens (human)   |   |       |       |
| ORGANISM                  | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. |   |       |       |
| REFERENCE                 | 1 (bases 1 to 734)   |   |       |       |

|                           |   |   |       |             |
|---------------------------|---|---|-------|-------------|
| AUTHORS                   | NIH-MGC http://mgc.nci.nih.gov/.  |   |       |             |
| TITLE                     | National Institutes of Health, Mammalian Gene Collection (MGC)  |   |       |             |
| JOURNAL                   | Unpublished   |   |       |             |
| COMMENT                   | Contact: Robert Strausberg, Ph.D.<br>Email: cnapbs-r@mail.nih.gov<br>Tissue Procurement: ATCC<br>cDNA Library Preparation: Life Technologies, Inc.<br>cDNA Library Arrayed by: Incyte Genomics, Inc.<br>DNA Sequencing by: Incyte Genomics, Inc.<br>Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:<br>http://image.llnl.gov<br>Plate: L1AM11273 row: j column: 21<br>High quality sequence stop: 734. |   |       |             |
| FEATURES                  |   |   |       |             |
| source                    | 1..734  |   |       |             |
|                           | Location/Qualifiers   | 127 a   | 228 c | 253 g 126 t |
|                           | /organism="Homo sapiens"  |   |       |             |
|                           | /mol_type="mRNA"  |   |       |             |
|                           | /db_xref="taxon:9606"   |   |       |             |
|                           | /clone="IMAGE:5112692"  |   |       |             |
|                           | /tissue_type="cervical carcinoma cell line"   |   |       |             |
|                           | /lab_host="DH10B"   |   |       |             |
|                           | /clone_lib="NIH_MGC_12"   |   |       |             |
|                           | /note="Organ: cervix; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dr. Average insert size 1.4 kb. Library prepared by Life Technologies."   |   |       |             |
| BASE COUNT                | 127 a   | 228 c   | 253 g | 126 t       |
| ORIGIN                    |   |   |       |             |
| Query Match               | 51.4%; Score 588.2; DB 12; Length 734;  |   |       |             |
| Best Local Similarity     | 98.6%; Pred. No. 2.6e-129;  |   |       |             |
| Matches 715; Conservative | 0; Mismatches 8; Indels 2; Gaps 2;  |   |       |             |
| Qy                        | 36  | GCTCGGAGCGCTCCGCGCGGTGGCTTCGTGGCGCGGCCCGCATGGCGGGGCC        | 95    |             |
| Db                        | 11  | GCTCGGAGCGCTCCGCGCGGTGGGGTCTGGCGCGGCCCGCATGGCGGGGCC         | 70    |             |
| Qy                        | 96  | ATCGCGAGGGCCCTCATCAGACGAGGAAAAGTGAAGCTCAGCACATCTAGGCAGTGCA  | 155   |             |
| Db                        | 71  | ATCGCGAGGGCCCTCATCAGACGAGGAAAAGTGAAGCTCAGCACATCTAGGCAGTGCA  | 130   |             |
| Qy                        | 156   | CCAACACAGGAACCTATGTCACCTTCAAGCTCTGGGTGGCGGACCGCAGCTCCAC     | 215   |             |
| Db                        | 131   | CCAACACAGGAACCTATGTCACCTTCAAGCTCTGGGTGGCGGACCGCAGCTCCAC     | 190   |             |
| Qy                        | 216   | CAGGAGTGTGTGACAGCTGCTGCTCATCTTTGCCACCAAGCCTCATGTGCTGCCA     | 275   |             |
| Db                        | 191   | CAGGAGTGTGTGACAGCTGCTGCTCATCTTTGCCACCAAGCCTCATGTGCTGCCA     | 250   |             |
| Qy                        | 276   | GCTGTCTGGGAGGTGGCTTCTGTGGTCAACACTGAACACATCTTTGGTCCGTGGCT    | 335   |             |
| Db                        | 251   | GCTGTCTGGGAGGTGGCTTCTGTGGTCAACACTGAACACATCTTTGGTCCGTGGCT    | 310   |             |
| Qy                        | 336   | GCTGGGTGTCTCTGACACCCCTGGAGGAGCTGCTGCCCCCAACACACAGGCTGTCGG   | 395   |             |
| Db                        | 311   | GCTGGGTGTCTCTGACACCCCTGGAGGAGCTGCTGCCCCCAACACACAGGCTGTCGG   | 370   |             |
| Qy                        | 396   | GTCTTGCCCAACCTGCCCTGTGTGTCAGAGGGGCCATAGTGTGGCGCGGGGCCG      | 455   |             |
| Db                        | 371   | GTCTTGCCCAACCTGCCCTGTGTGTCAGAGGGGCCATAGTGTGGCGCGGGGCCG      | 430   |             |
| Qy                        | 456   | CACGTGGGAGCAGCAGAACAGCTCCTGCAGCATCTGCTGGAGGCCCTGTGGCGGTGT   | 515   |             |
| Db                        | 431   | CACGTGGGAGCAGCAGAACAGCTCCTGCAGCATCTGCTGGAGGCCCTGTGGCGGTGT   | 490   |             |
| Qy                        | 516   | GAGGAGTGTCTGAAGCCTAGCTGCAGATCCACACTGGCCTCAGTGGCAGTGGCGGCC   | 575   |             |
| Db                        | 491   | GAGGAGTGTCTGAAGCCTAGCTGCAGATCCACACTGGCCTCAGTGGCAGTGGCGGCC   | 550   |             |
| Qy                        | 576   | TTCTGTGTGCATTTCTCCGAGGCCCTGGCTGAAGGAGCCGTCAAGATGGGCATGCCAGC | 635   |             |
| Db                        | 551   | TTCTGTGTGCATTTCTCCGAGGCCCTGGCTGAAGGAGCCGTCAAGATGGGCATGCCAGC | 610   |             |

QY 636 AGCCTGGCCACCCGATCGCTGCCAGACCCCTGCTGGGAGGCGCAAGATGCTGTCAC 695  
|||||  
Db 611 AGCCTGGCCACCCGATCGCTGCCAGACCCCTGCTGGGAGGCGCAAGATGCTGTCAC 669  
|||||  
QY 696 G-AGGCCAACAACCCAGCCAGCTGGCTCAGACCGTGTGCACCCCGGTGGCAGCACCAT 754  
|||||  
Db 670 GAAGGCCAAAACCCAGCCAGCTGGCTCAGACCGTGTGCACCCCGGTGGCAGCACCAT 729  
|||||  
QY 755 CTATG 759  
|||||  
Db 730 CTATG 734

Search completed: August 8, 2003, 17:47:43  
Job time : 2686 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 8, 2003, 14:36:31 ; Search time 4867 Seconds  
(without alignments)  
11246.575 Million cell updates/sec

Title: US-09-806-536A-29  
Perfect score: 1338  
Sequence: 1 ggttagcgcagctctgtccga.....ggggcagcagcatatgggggt 1338

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl.\*

- 1: gb.ba.\*
- 2: gb.htg.\*
- 3: gb.in.\*
- 4: gb.om.\*
- 5: gb.ov.\*
- 6: gb.pat.\*
- 7: gb.ph.\*
- 8: gb.pl.\*
- 9: gb.pr.\*
- 10: gb.ro.\*
- 11: gb.sts.\*
- 12: gb.sy.\*
- 13: gb.un.\*
- 14: gb.vi.\*
- 15: em.ba.\*
- 16: em.fun.\*
- 17: em.hum.\*
- 18: em.in.\*
- 19: em.mu.\*
- 20: em.om.\*
- 21: em.ov.\*
- 22: em.pat.\*
- 23: em.ph.\*
- 24: em.pl.\*
- 25: em.ro.\*
- 26: em.sts.\*
- 27: em.un.\*
- 28: em.vi.\*
- 29: em.htg\_hum.\*
- 30: em.htg\_inv.\*
- 31: em.htg\_other.\*
- 32: em.htg\_mus.\*
- 33: em.htg\_pin.\*
- 34: em.htg\_rod.\*
- 35: em.htg\_mam.\*
- 36: em.htg\_vrt.\*
- 37: em.sy.\*
- 38: em.htgo\_hum.\*
- 39: em.htgo\_mus.\*
- 40: em.htgo\_other.\*
- 41: em.htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score  | Query Match % | Length | DB | ID        | Description         |
|------------|--------|---------------|--------|----|-----------|---------------------|
| 1          | 1307   | 97.7          | 2331   | 6  | BD160280  | BD160280 Primer fo  |
| 2          | 1307   | 97.7          | 2331   | 9  | AK023914  | AK023914 Homo sapi  |
| c          | 1291.8 | 96.5          | 2615   | 6  | BD156362  | BD156362 Primer fo  |
| c          | 1291.8 | 96.5          | 2615   | 9  | AK001500  | AK001500 Homo sapi  |
| 5          | 1200.4 | 89.7          | 2400   | 9  | HS0805145 | AL833857 Homo sapi  |
| 6          | 1127.8 | 84.3          | 1178   | 9  | BC007993  | BC007993 Homo sapi  |
| 7          | 737.6  | 55.1          | 794    | 6  | BD150492  | BD150492 Primer fo  |
| c          | 653.2  | 48.8          | 101777 | 9  | AC105210  | AC105210 Homo sapi  |
| c          | 653.2  | 48.8          | 184959 | 2  | AC087823  | AC087823 Homo sapi  |
| c          | 653.2  | 48.8          | 174906 | 9  | AC067930  | AC067930 Homo sapi  |
| 11         | 599.4  | 44.8          | 207636 | 2  | AC019122  | AC019122 Homo sapi  |
| 12         | 567.6  | 42.4          | 1276   | 10 | BC026536  | BC026536 Mus muscu  |
| 13         | 486.6  | 36.4          | 563    | 6  | BD148256  | BD148256 Primer fo  |
| c          | 338.4  | 25.3          | 197243 | 2  | AC120593  | AC120593 Rattus no  |
| 15         | 226.4  | 16.9          | 1848   | 9  | BC001504  | BC001504 Homo sapi  |
| 16         | 224.8  | 16.8          | 1792   | 6  | AX409754  | AX409754 Sequence   |
| 17         | 224.8  | 16.8          | 1792   | 9  | HUMP5CR   | M77836 Human pyro   |
| 18         | 213.8  | 16.0          | 1769   | 9  | AF218000  | AF218000 Homo sapi  |
| 19         | 205.4  | 15.4          | 1757   | 9  | BC022244  | BC022244 Homo sapi  |
| 20         | 201.8  | 15.1          | 1676   | 9  | BC014868  | BC014868 Homo sapi  |
| 21         | 201.8  | 15.1          | 1708   | 9  | BC020553  | BC020553 Homo sapi  |
| 22         | 191.6  | 14.3          | 1590   | 10 | BC006882  | BC006882 Mus muscu  |
| 23         | 173.4  | 13.0          | 1820   | 10 | BC006727  | BC006727 Mus muscu  |
| 24         | 167    | 12.5          | 999    | 9  | AF151351  | AF151351 Homo sapi  |
| 25         | 156.4  | 11.7          | 299800 | 1  | AP005040  | AP005040 Streptomy  |
| c          | 151.6  | 11.3          | 293050 | 1  | SCO939116 | AL939116 Streptomy  |
| 27         | 148.8  | 11.1          | 3149   | 6  | BD156370  | BD156370 Primer fo  |
| 28         | 148.8  | 11.1          | 3149   | 9  | AK001508  | AK001508 Homo sapi  |
| c          | 147    | 11.0          | 148283 | 2  | AC0116520 | AC0116520 Mus muscu |
| c          | 143.4  | 10.7          | 10490  | 1  | AE013306  | AE013306 Methanosa  |
| 31         | 143.2  | 10.7          | 178292 | 2  | AC120830  | AC120830 Rattus no  |
| 32         | 143.2  | 10.7          | 245651 | 2  | AC126537  | AC126537 Rattus no  |
| 33         | 141.4  | 10.6          | 999    | 3  | AY071558  | AY071558 Drosophil  |
| 34         | 138.4  | 10.3          | 822    | 1  | AY142817  | AY142817 Heliohaci  |
| 35         | 138.4  | 10.3          | 5139   | 1  | AF305580  | AF305580 Methanosa  |
| 36         | 138.4  | 10.3          | 10357  | 1  | AE011122  | AE011122 Methanosa  |
| c          | 127.6  | 9.5           | 28128  | 2  | AC015301  | AC015301 Drosophil  |
| 38         | 127.6  | 9.5           | 171612 | 3  | AC009461  | AC009461 Drosophil  |
| 39         | 127.6  | 9.5           | 175440 | 3  | AC007804  | AC007804 Drosophil  |
| 40         | 127.6  | 9.5           | 253101 | 3  | AE003718  | AE003718 Drosophil  |
| 41         | 126    | 9.4           | 10029  | 1  | AE012787  | AE012787 Chlorobiu  |
| 42         | 123    | 9.2           | 2678   | 1  | TTTHPROC  | D25413 Thermus the  |
| 43         | 118.2  | 8.8           | 309350 | 1  | AP005215  | AP005215 Corynebac  |
| 44         | 116.6  | 8.7           | 1089   | 3  | AF098020  | AF098020 Drosophil  |
| 45         | 111.8  | 8.4           | 4290   | 3  | AF170829  | AF170829 Drosophil  |

ALIGNMENTS

RESULT 1  
BD160280  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE

BD160280  
Primer for synthesizing full-length cDNA and use thereof.  
BD160280  
BD160280.1 GI:27866038  
JP 2002191363-A/15123.  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 2331)  
Ota.T., Isogai.T., Nishikawa.T., Hayashi.K., Saito.K., Yamamoto.J.,  
Ishii.S., Sugiyama.T., Wakamatsu.A., Nagai.K. and Otsuki.T.  
Primer for synthesizing full-length cDNA and use thereof

2331 bp DNA linear PAT 17-JAN-2003

JOURNAL

Patent: JP 2002191363-A 15123 09-JUL-2002;

COMMENT

HELIX RESEARCH INSTITUTE  
OS Homo sapiens (human)  
PN JP 2002191363-A/15123  
PD 09-JUL-2002  
PF 28-JUL-2000 JP 2000280990  
PI TOSHIO OTA,TAKAO ISOGAI,TETSUO NISHIKAWA,KOJI HAYASHI,KAORU  
PI SAITO,  
PI JUNICHI YAMAGATA,SHIZUKO ISHII,TOMOYASU SUGIYAMA,AI WAKAMATSU,  
PI KEIICHI NAGAI,TETSUJI OTSUKI  
PC C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/PC  
10,  
PC C12P21/02,C12P21/08,C06F17/30,C12N15/00,C12N5/00 CC  
Primer for synthesizing full-length cDNA and use thereof FH Key  
Location/Qualifiers  
FT CDS  
Location/Qualifiers  
1..2331  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
BASE COUNT 398 a 753 c 748 g 432 t  
ORIGIN

FEATURES  
source

Query Match 97.7%; Score 1307; DB 6; Length 2331;  
Best Local Similarity 99.2%; Pred. No. 1e-224;  
Matches 1313; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 15 GTCCGAGGCAACAAAGATGGCAGCTGCGGAGCGCTCTCCGGCGCGGTGGGCTTCGTGGGC 74  
Db 12 GTCCGAGGCAACAAAGATGGCAGCTGCGGAGCGCTCTCCGGCGCGGTGGGCTTCGTGGGC 71  
Qy 75 GCGGGCGCATGTCGGGGGCCATTCGCGCAGGGCGCTCATCAGAGCAGAGAAAGTGAAGCT 134  
Db 72 GCGGGCGCATGTCGGGGGCCATTCGCGCAGGGCGCTCATCAGAGCAGAGAAAGTGAAGCT 131  
Qy 135 CAGCACATCTGCGCAGTGCACCAACAGACAGAACTATGTCACTTTCAAGCTCTGGGT 194  
Db 132 CAGCACATCTGCGCAGTGCACCAACAGACAGAACTATGTCACTTTCAAGCTCTGGGT 191  
Qy 195 TGCCGGACACGACCTCCACACAGGAGGTGCTGCAGAGCTGCTGCTGCTCATCTTTGCC 254  
Db 192 TGCCGGACACGACCTCCACACAGGAGGTGCTGCAGAGCTGCTGCTGCTCATCTTTGCC 251  
Qy 255 ACCAAGCCTCATGTGCTGCAGAGTGTCTGGCAGAGGTGGCTCTGTGTCCACCACTGAA 314  
Db 252 ACCAAGCCTCATGTGCTGCAGAGTGTCTGGCAGAGGTGGCTCTGTGTCCACCACTGAA 311  
Qy 315 CACATCTTGGTGTCCGTGGCTGCTGGGGTGTCTCTGAGCACCTGGAGAGCTGTGCC 374  
Db 312 CACATCTTGGTGTCCGTGGCTGCTGGGGTGTCTCTGAGCACCTGGAGAGCTGTGCC 371  
Qy 375 CCAACACACGGGTGCTGGGGTCTTGGCCACCTGCTGTGTGTGGTCCAGGAGGGCC 434  
Db 372 CCAACACACGGGTGCTGGGGTCTTGGCCACCTGCTGTGTGTGGTCCAGGAGGGCC 431  
Qy 435 ATAGTGTATGGCGGGGGCCGACCTGGGGACCGAGACCAAGCTCCTCGAGCATGTG 494  
Db 432 ATAGTGTATGGCGGGGGCCGACCTGGGGACCGAGACCAAGCTCCTCGAGCATGTG 491  
Qy 495 CTGGAGGCGCTGTGGGCGGTGTGAGGAGGTGCTGAAGCTACGTGACATCCACACTGGC 554  
Db 492 CTGGAGGCGCTGTGGGCGGTGTGAGGAGGTGCTGAAGCTACGTGACATCCACACTGGC 551  
Qy 555 CTCAGTGGCAGTGGCGGTGCTTGTGTGTGCACTTCCTCGAGGCGCTGGCTGAAGAGGCC 614  
Db 552 CTCAGTGGCAGTGGCGGTGCTTGTGTGTGCACTTCCTCGAGGCGCTGGCTGAAGAGGCC 611  
Qy 615 GTCAAGATGGGCATGCCAGCAGCTGGCCACCGCATCGTGGCCAGACCTCTGCTGGG 674  
Db 612 GTCAAGATGGGCATGCCAGCAGCTGGCCACCGCATCGTGGCCAGACCTCTGCTGGG 671

Qy 675 ACGGCCAAAGATGCTGCTGCACGAGGGCCAAACACCAGCCAGCCAGCTGCGCTCAGAGCTGTGC 734  
Db 672 ACGGCCAAAGATGCTGCTGCACGAGGGCCAAACACCAGCCAGCTGCGCTCAGAGCTGTGC 731  
Qy 735 ACCCGGGTGGCACCACCATCTATGACTTCCACGCCCTGGAGCAGGGCGGCTCGGAGCA 794  
Db 732 ACCCGGGTGGCACCACCATCTATGACTTCCACGCCCTGGAGCAGGGCGGCTCGGAGCA 791  
Qy 795 GCCACCATAGCGCGCTGGAGGCTGCCACCTGCCGGGCCAAAGAGCTCAGCAGAAAGTAG 854  
Db 792 GCCACCATAGCGCGCTGGAGGCTGCCACCTGCCGGGCCAAAGAGCTCAGCAGAAAGTAG 851  
Qy 855 GCTGGGCTCTGCGCATCTTCCCTGCTGTGCCCTGTGTGCCCTGTCCCTGTGTGCTTCCC 914  
Db 852 GCTGGGCTCTGCGCATCTTCCCTGCTGTGCCCTGTGTGCCCTGTCCCTGTGTGCTTCCC 911  
Qy 915 CTGAGGACTGGGGCTCCCTCCTCTGTCATGAGGCTCTCCTACTGCTCTCTTCTCCCTTTG 974  
Db 912 CTGAGGACTGGGGCTCCCTCCTCTGTCATGAGGCTCTCCTACTGCTCTCTTCTCCCTTTG 971  
Qy 975 CACAGGGAATGACAGGGGGCAGGACTTGGGAGGTTCACAGCAGGGGGAGCCGCCAGCA 1034  
Db 972 CACAGGGAATGACAGGGGGCAGGACTTGGGAGGTTCACAGCAGGGGGAGCCGCCAGCA 1031  
Qy 1035 GTGGGACACTCTCCCTCCCTCAGTGAGCAGAGGAGCCGCTGGTGGCTCTGCGCCCTT 1094  
Db 1032 GTGGGACACTCTCCCTCCCTCAGTGAGCAGAGGAGCCGCTGGTGGCTCTGCGCCCTT 1091  
Qy 1095 GCTGAGTGAGCCCACTTGTGCAACATGTTCTGAGGGGCCCAAGAGATGGCGTCTT 1154  
Db 1092 GCTGAGTGAGCCCACTTGTGCAACATGTTCTGAGGGGCCCAAGAGATGGCGTCTT 1151  
Qy 1155 GGTCAATTTGCCCGCATGGTTGGGAGTGGTGTGAGGCCATGAACAGAACTTACGGTAACA 1214  
Db 1152 GGTCAATTTGCCCGCATGGTTGGGAGTGGTGTGAGGCCATGAACAGAACTTACGGTAACA 1211  
Qy 1215 GGCACGGCTGGCCCAATGCTGCTGGAGCTGGAGCTTGGCTTTGGCTTTTCAAGTGGG 1274  
Db 1212 GGCACGGCTGGCCCAATGCTGCTGGAGCTGGAGCTTGGCTTTGGCTTTTCAAGTGGG 1271  
Qy 1275 CTCGTGAGCTACAGCCAGCGCGCTGCCCTCATCTCAGTCTAGGGGGCAGCAGCATATG 1334  
Db 1272 TCCGTGAGCTACAGCCAGCGCGCTGCCCTCATCTCAGTCTAGGGGGCAGCAGCATAT 1331  
Qy 1335 GGG 1337  
Db 1332 GGG 1334

RESULT 2

AK023914

LOCUS

DEFINITION

Homo sapiens cDNA FLJ13852 fis, clone THYR0100934, weakly similar

to PTEROLINE-5-CARBOXYLATE REDUCTASE (EC 1.5.1.2).

AK023914

ACCESSION

VERSION

AK023914.1 GI:10435995

KEYWORDS

oligo capping; fis (full insert sequence).

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1

Isogai, T., Ota, T., Hayashi, K., Sugiyama, T., Otsuki, T., Suzuki, Y.,

Nishikawa, T., Nagai, K., Sugano, S., Takahashi-Fujii, A., Hara, H.,

Tanase, T., Nomura, Y., Togiya, S., Komai, F., Hara, R., Takeuchi, K.,

Arita, M., Nabekura, T., Ishii, S., Kawai, Y., Saito, K., Yamamoto, J.,

Wakamatsu, A., Nakamura, Y., Nagaharai, K., Masuno, Y. and Oshima, A.

NEDO human cDNA sequencing project

Unpublished

2 (bases 1 to 2331)

Isogai, T. and Otsuki, T.

Direct Submission

TITLE

JOURNAL

Submitted (23-AUG-2000) Takao Isogai, Helix Research Institute,



Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan  
(E-mail: genomics@shri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)

## COMMENT

NEDO human cDNA sequencing project supported by Ministry of  
International Trade and Industry of Japan; cDNA full insert  
sequencing; Research Association for Biotechnology; cDNA library  
construction; 5'- & 3'-end one pass sequencing and clone selection;  
Helix Research Institute (supported by Japan Key Technology Center  
etc.) and Department of Virology, Institute of Medical Science,  
University of Tokyo.

## FEATURES

Location/Qualifiers

1..2331  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="THYRO1000934"  
/tissue\_type="thyroid gland"  
/clone\_lib="THYRO1"  
/note="cloning vector: pME18SFL3"  
27..851

## CDS

/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="BAB14721.1"  
/db\_xref="GI:10435996"  
/translation="MAAEPSPRRVFGVAGRMAGIAQGLIRAGKVAQHILASAPT  
DRNLCHFQALGCRTHNSQVLOSLVIFATKPHVLPVLAIVAPVVTTHILVSV  
AGVSLTLELLPPNTHRVLRVLPVQVQGAIVMGRHVGSSEKLLHLLFNC  
REEVPEAYVDIHTGLSGVAFVCFSEALAGAVKMGMESSLAHRIATAQLGTAK  
MLLHFGQHPAQLRSDVCTPGGTIYGLHLEPQGLRAATMSAVEAATRAKELSRK"  
BASE COUNT 398 a 753 c 748 g 432 t

Query Match 97.7%; Score 1307; DB 9; Length 2331;  
Best Local Similarity 99.2%; Pred. No. 1e-224; Indels 0; Gaps 0;  
Matches 1313; Conservative 0; Mismatches 10;

|    |     |  |     |
|----|-----|--|-----|
| QY | 15  | GTCCGAGGCAACAAGATGCGAGCTGCGGAGCGCTCCCGCGCGCGTTCGTGGGC        | 74  |
| DB | 12  | GTCCGAGGCAACAAGATGCGAGCTGCGGAGCGCTCCCGCGCGCGTTCGTGGGC        | 71  |
| QY | 75  | GGCGGCGGATGGCGGGGCCATCGCGAGGGCCCTATCAGAGAGGAAAGTGAAGCT       | 134 |
| DB | 72  | GGCGGCGGATGGCGGGGCCATCGCGAGGGCCCTATCAGAGAGGAAAGTGAAGCT       | 131 |
| QY | 135 | CAGCACAATGCGCCAGTGCACCAACAGACAGAACTATGTCATTTCAAGCTCTGGT      | 194 |
| DB | 132 | CAGCACAATGCGCCAGTGCACCAACAGACAGAACTATGTCATTTCAAGCTCTGGT      | 191 |
| QY | 195 | TGCGGACACACGCTCCCAACAGGAGGTGCTGCAGAGCTGCTCGTCATCTTTGCC       | 254 |
| DB | 192 | TGCGGACACACGCTCCCAACAGGAGGTGCTGCAGAGCTGCTCGTCATCTTTGCC       | 251 |
| QY | 255 | ACCAAGCCTATGTGTCGACGCTGCTCCGAGAGGTGCTGAGAGGCTGCTGCGCC        | 314 |
| DB | 252 | ACCAAGCCTATGTGTCGACGCTGCTCCGAGAGGTGCTGAGAGGCTGCTGCGCC        | 311 |
| QY | 315 | CACATCTGTGTCGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG      | 374 |
| DB | 312 | CACATCTGTGTCGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG      | 371 |
| QY | 375 | CCAAACACACGCGTGTGGGGTCTTGGCCCAACCTGCTGCTGCTGCTGCTGCTGCTG     | 434 |
| DB | 372 | CCAAACACACGCGTGTGGGGTCTTGGCCCAACCTGCTGCTGCTGCTGCTGCTGCTG     | 431 |
| QY | 435 | ATAGTGATGGCGGGGGCGCCACGCTGGGAGCAGGAGCAAGCTCTCGCAGCATCTG      | 494 |
| DB | 432 | ATAGTGATGGCGGGGGCGCCACGCTGGGAGCAGGAGCAAGCTCTCGCAGCATCTG      | 491 |
| QY | 495 | CTGAGGCGCTGTGGGGGTGTGAGAGGTGCTGAGGCTAGCTGAGCTCAGCATCCACACTGC | 554 |
| DB | 492 | CTGAGGCGCTGTGGGGGTGTGAGAGGTGCTGAGGCTAGCTGAGCTCAGCATCCACTGC   | 551 |
| QY | 555 | CTCAGTGGCAGTGGCGTGGCTTCGTGTGTCATTTCCGAGGCGCTGGCTGAAGGAGCC    | 614 |

|    |      |   |      |
|----|------|---|------|
| Db | 552  | CTCAGTGGCAGTGGCGTGGCTTCGTGTGTGCAATTCCTCCGAGGCGCTGGCTGAAGGAGCC | 611  |
| QY | 615  | GTCAAGATGGCATGGCCAGCAGCCTGGCCACCGCATCGCTGCCAGAGCCCTGCTGGGG    | 674  |
| Db | 612  | GTCAAGATGGCATGGCCAGCAGCCTGGCCACCGCATCGCTGCCAGAGCCCTGCTGGGG    | 671  |
| QY | 675  | ACGCCCAAGATGCTGCTGTCAGCAGGCGCAACACCCAGCCAGCTGCGCTCAGAGCTGTC   | 734  |
| Db | 672  | ACGCCCAAGATGCTGCTGTCAGCAGGCGCAACACCCAGCCAGCTGCGCTCAGAGCTGTC   | 731  |
| QY | 735  | ACCCCGGTGGCAGCAGCCTATGAGACTACAGCCCTGAGCAGAGGCGGGCTGCGGAGCA    | 794  |
| Db | 732  | ACCCCGGTGGCAGCAGCCTATGAGACTACAGCCCTGAGCAGAGGCGGGCTGCGGAGCA    | 791  |
| QY | 795  | GCCACCATGAGCGCGTGGAGGTGCCACCTGCGCGGCCAAGAGCTCAGCAGAAAGTAG     | 854  |
| Db | 792  | GCCACCATGAGCGCGTGGAGGTGCCACCTGCGCGGCCAAGAGCTCAGCAGAAAGTAG     | 851  |
| QY | 855  | GCTGGGCTCTGGCCATCTTCTCCCTCTGTGCGCCCTGCTCCCTGTGCTGCTCCCTTCCC   | 914  |
| Db | 852  | GCTGGGCTCTGGCCATCTTCTCCCTCTGTGCGCCCTGCTCCCTGTGCTGCTCCCTTCCC   | 911  |
| QY | 915  | CTGAGGACTGCGGCTCCCTCCTCTGATGAGGCTCTCTCTGCTGCTCTCTCTCTCTCTG    | 974  |
| Db | 912  | CTGAGGACTGCGGCTCCCTCCTCTGATGAGGCTCTCTCTGCTGCTCTCTCTCTCTCTG    | 971  |
| QY | 975  | CACAGGAAATGACAGGGGCGAGACTTGGGAGGTTCAGCAGGCGGGGAGCCCGGAGCA     | 1034 |
| Db | 972  | CACAGGAAATGACAGGGGCGAGACTTGGGAGGTTCAGCAGGCGGGGAGCCCGGAGCA     | 1031 |
| QY | 1035 | GTGGGGACACTCTCTCCCTCCCTGAGCAGCAAGGACCGTGGTGGTCTCTGCTCTCTCT    | 1094 |
| Db | 1032 | GTGGGGACACTCTCTCCCTCCCTGAGCAGCAAGGACCGTGGTGGTCTCTGCTCTCTCT    | 1091 |
| QY | 1095 | GCTCAGTGGAGCCACCTTGTGCAACATTGGTTCTGAGGGGCCCAAGAGATGCGCTTT     | 1154 |
| Db | 1092 | GCTCAGTGGAGCCACCTTGTGCAACATTGGTTCTGAGGGGCCCAAGAGATGCGCTTT     | 1151 |
| QY | 1155 | GGTCATTTGGCCGATGTTGGCAGTTGGTTGAGGCGCATGACAGCAACTTACCGTAACA    | 1214 |
| Db | 1152 | GGTCATTTGGCCGATGTTGGCAGTTGGTTGAGGCGCATGACAGCAACTTACCGTAACA    | 1211 |
| QY | 1215 | GGCAGGCTGCCCAATGCTTGGTCTGAGCTGGAGCTGCTTTTGGCTTTTCCAGTGGG      | 1274 |
| Db | 1212 | GGCAGGCTGCCCAATGCTTGGTCTGAGCTGGAGCTGCTTTTGGCTTTTCCAGTGGG      | 1271 |
| QY | 1275 | CTCGTGAGCTACAGCGCGGCTGCTCATCTCAGCTCTAGGGGCGCAGCAGCATATG       | 1334 |
| Db | 1272 | TCCGTGAGCTACAGCGCGGCTGCTCATCTCAGCTCTAGGGGCGCAGCAGCATAT        | 1331 |
| QY | 1335 | GGG 1337  |      |
| Db | 1332 | GGG 1334  |      |

## RESULT 3

BD156362/c

LOCUS

BD156362

DEFINITION

Primer for synthesizing full-length cDNA and use thereof.

ACCESSION

BD156362.1

VERSION

BD156362.1

KEYWORDS

JP 2002191363-A/11205.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 2615)

AUTHORS

Ota, T., Isogai, T., Nishikawa, T., Hayashi, K., Saito, K., Yamamoto, J.,

Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K. and Otsuki, T.

Primer for synthesizing full-length cDNA and use thereof

Patent: JP 2002191363-A 11205 09-JUL-2002;

HELIIX RESEARCH INSTITUTE

OS Homo sapiens (human)



[illegible]

|                            |   |
|----------------------------|---|
| gene                       | 1. .2400  |
| CDS                        | /gene="DKFzp761H0716"<br><1..735<br>/note="similarity to pyrroline 5-carboxylate reductase isoform (Homo sapiens)"<br>/codon_start=1<br>/product="hypothetical protein"<br>/protein_id="CAD38716.1"<br>/db_xref="GI:21739341"<br>/translation="CKVEAQHTLASAPTDRLNCHFOALGCRTHSNQEVLOSCLLVIFAKPHVLPAVLAIEAVPVTEHLIVSAGVSLSLELLPPNTRVLKRVLPNPFCVVQEGAINARGRHVSGSETNILLOHLEACRCREVPYAYDIHTGLSGSVAFVCFASFLAEAGVMGPSSLAHRTAAOPLLTAKMLLHEQHQAQLRSDVCTPGGTIIYLHALEOGLRAATMSVAEATCAKELSRK" |
| polyA_site                 | 2381  |
| BASE COUNT                 | 452 a 756 c 740 g 452 t   |
| ORIGIN                     |   |
| Query Match                | 89.7%; Score 1200.4; DB 9; Length 2400;   |
| Best Local Similarity      | 99.1%; Pred. No. 1.3e-205;  |
| Matches 1207; Conservative | 0; Mismatches 11; Indels 0; Gaps 0;   |
| Qy                         | 120 GGAAAAGTGAAGCTCAGCACATACTGGCCAGTCGCCAACACAGACAGAACTATGTCAC 179  |
| Db                         | 1 GGAAAAGTGAAGCTCAGCACATACTGGCCAGTCGCCAACACAGACAGAACTATGTCAC 60   |
| Qy                         | 180 TTTCAAGCTCTGGTTGCCCGACCAGGCATCTCAACAGAGAGTGCTGCAGAGTGCCTG 239   |
| Db                         | 61 TTTCAAGCTCTGGTTGCCCGACCAGGCATCTCAACAGAGAGTGCTGCAGAGTGCCTG 120  |
| Qy                         | 240 CTCGTCATCTTTGCCACCAAGCCTCATGTGCTGCCAGCTGTCTGGCAGAGGTGSCCT 299   |
| Db                         | 121 CTCGTCATCTTTGCCACCAAGCCTCATGTGCTGCCAGCTGTCTGGCAGAGGTGGCTCT 180  |
| Qy                         | 300 GTGGTCACCACTGAACACATCTTGGTGCTCGGTGCTGGGGTGCTCTGAGCACCGTG 359  |
| Db                         | 181 GTGGTCACCACTGAACACATCTTGGTGCTCGGTGCTGGGGTGCTCTGAGCACCGTG 240  |
| Qy                         | 360 GAGGAGCTGCTGCCCCCAAACACAGGGTGCTCGGGGTCTGGCCAACTGCCCTGTGTG 419   |
| Db                         | 241 GAGGAGCTGCTGCCCCCAAACACAGGGTGCTCGGGGTCTGGCCAACTGCCCTGTGTG 300   |
| Qy                         | 420 GTCCAGGAAGGGGCCATAGTGATGGCGGGGCCACACGTGGGAGCAGGACCAAG 479   |
| Db                         | 301 GTCCAGGAAGGGGCCATAGTGATGGCGGGGCCACACGTGGGAGCAGGACCAAC 360   |
| Qy                         | 480 CTCTGCAGCATCTGCTGGAGCCCTGTGGCGGTGTGAGAGGTGCCTGAAGCCTACGTC 539   |
| Db                         | 361 CTCTGCAGCATCTGCTGGAGCCCTGTGGCGGTGTGAGAGGTGCCTGAAGCCTACGTC 420   |
| Qy                         | 540 GACATCCACTGGCTCAGTGGCAGTGGCGTTCGTTGTGTGANTCTCCGAGGCC 599  |
| Db                         | 421 GACATCCACTGGCTCAGTGGCAGTGGCGTTCGTTGTGTGANTCTCCGAGGCC 480  |
| Qy                         | 600 CTGGCTGAAGGAGCCGCTCAAGATGGCATGCCACAGCCCTGGCCCAGCATCGCTGCC 659   |
| Db                         | 481 CTGGCTGAAGGAGCCGCTCAAGATGGCATGCCACAGCCCTGGCCCAGCATCGCTGCC 540   |
| Qy                         | 660 CAGACCTCTGCTGGGACGCCCAAGATGCTGTGCACAGGGGCCAACCCAGCCACGCTG 719   |
| Db                         | 541 CAGACCTCTGCTGGGACGCCCAAGATGCTGTGCACAGGGGCCAACCCAGCCACGCTG 600   |
| Qy                         | 720 CGCTCAGAGCTGTGCACCCCGGGTGGCACACCACTATTATNGACTTCCAGCCCTGGAGCAG 779   |
| Db                         | 601 CGCTCAGAGCTGTGCACCCCGGGTGGCACACCACTATTATNGACTTCCAGCCCTGGAGCAG 660   |
| Qy                         | 780 GCGGGCTCGGAGCAGCCACCATGAGCCCTGGAGGCTGCCACCTCCGGGCCAAGGAG 839  |
| Db                         | 661 GCGGGCTCGGAGCAGCCACCATGAGCCCTGGAGGCTGCCACCTCCGGGGCCAAGGAG 720   |
| Qy                         | 840 CTCAGCAGAAGTAGGCTGGGCTCTGGGCATCTCTTCTGCTGCTGCTGCTGCTCTC 899   |

|            |  |  |  |
|------------|--|--|--|
| Db         | 721  | CTCAGCAGAAAGTAGGCTGGGCTCTGGCCATCCTTTCTGCTGCTGTGCCCCCTGCCTCTC | 780  |
| Qy         | 900  | CCTGTGTCCCTTCCCTCAGGACTCGCGCTCCCTCCCTGCGATGAGGCTCTCCCTACTG   | 959  |
| Db         | 781  | CCTGTGTCCCTTCCCTCAGGACTCGCGCTCCCTCCCTGCGATGAGGCTCTCCCTACTG   | 840  |
| Qy         | 960  | CTCCTTCTCCCTTGCACAGGGGAAATGCAGGGGGCAGGACTTGGGAGGTTCCACGAGCG  | 1019   |
| Db         | 841  | CTCCTTCTCCCTTGCACAGGGGAAATGCAGGGGGCAGGACTTGGGAGGTTCCACGAGCG  | 900  |
| Qy         | 1020   | GCGGAGCCCGCAGCAGTGGGACACATCCTCCTCCCGAGTGAGCAGAAAGCACCCTGGTG  | 1079   |
| Db         | 901  | GCGGAGCCCGCAGCAGTGGGACACATCCTCCTCCCGAGTGAGCAGAAAGCACCCTGGTG  | 960  |
| Qy         | 1080   | GTGGCTCTGCCCTTGTGTCAGTGAGCCACCTTGTGTCACACATTTGTTCTGAGGGGCC   | 1139   |
| Db         | 961  | GTGGCTCTGCCCTTGTGTCAGTGAGCCACCTTGTGTCACACATTTGTTCTGAGGGGCC   | 1020   |
| Qy         | 1140   | AAGAGATGGCGTCTTGGTTCATTTGCCCGCATGTTGGCAGTGTGGTTCAGCCATGAACA  | 1199   |
| Db         | 1021   | AAGAGATGGCGTCTTGGTTCATTTGCCCGCATGTTGGCAGTGTGGTTCAGCCATGAACA  | 1080   |
| Qy         | 1200   | GAACCTACGGTAACAGGCACGGCTGGCCCAATGCTGTCTGGAGCTTGGAGCTTTCCTTT  | 1259   |
| Db         | 1081   | GAACCTACGGTAACAGGCACGGCTGGCCCAATGCTGTCTGGAGCTTGGAGCTTTCCTTT  | 1140   |
| Qy         | 1260   | GGCTTTCAAGTGGGCTCGTGCAGCTACAGCCAGCGCGCTGCTCATCTCAGCTCTAGG    | 1319   |
| Db         | 1141   | GGCTTTCAAGTGGGCTCGTGCAGCTACAGCCAGCGCGCTGCTCATCTCAGCTCTAGG    | 1200   |
| Qy         | 1320   | GGGCACAGCATATGGGG 1337                                       |  |
| Db         | 1201   | GGGCACAGCATATGGG 1218  |  |
| RESULT 6   |  |  |  |
| BC007993   |  |  |  |
| LOCUS      |  |  |  |
| DEFINITION | Homo sapiens, hypothetical protein FLJ13852, clone MGC:14985   |  | PRI 12-JUL-2001  |
| IMAGE:     | 3533609, mRNA, complete cds.   |  |  |
| ACCESSION  | BC007993   |  |  |
| VERSION    | BC007993.1   |  |  |
| KEYWORDS   | MGC.   |  |  |
| SOURCE     | Homo sapiens (human)   |  |  |
| ORGANISM   | Homo sapiens   |  |  |
| REFERENCE  | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.   |  |  |
| AUTHORS    | Strausberg, R.   |  |  |
| TITLE      | Direct Submission  |  |  |
| JOURNAL    | Submitted (15-MAY-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA   |  |  |
| REMARK     | NIH-MGC Project URL: http://mgc.nci.nih.gov  |  |  |
| COMMENT    | Contact: MGC help desk<br>Email: cgapbs-re@mail.nih.gov<br>Tissue Procurement: DCTD/DTP<br>cdNA Library Preparation: Rubin Laboratory<br>CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)<br>DNA Sequencing by: Genome Sequence Centre,<br>BC Cancer Agency, Vancouver, BC, Canada<br>info@bcsc.bc.ca |  |  |
|            |  |  | Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saedi, Jacqueline Schein, Duane Smalish, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra. |

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAL Plate: 23 Row: i Column: 8  
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 10435995.

## FEATURES

source

Location/Qualifiers

1..1178

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:65263"

/db\_xref="taxon:9606"

/clone="MGC:14985 IMAGE:3533609"

/tissue\_type="Lung, small cell carcinoma"

/clone\_lib="NIH\_MGC\_7"

/lab\_host="DH10B-R"

/note="Vector: pOTB7"

47..871

/codon\_start=1

/product="hypothetical protein FLJ13852"

/protein\_id="AAH07993.1"

/db\_xref="GI:14124940"

/translation="MAAAPSPRRYFVGARMAGIAQGLIRAKVBAQHILASPT  
DRNLCHFALGRTHSNQEVLSCLVIFATKPHVLAFLAEPVVTTHILVSA  
AGVSLTELELPPTRLVRLVPLNCPVVOEGAIVMARGRHVGSSETNLLQLEACG  
RCVEEPAVDIHTGLSGGVAFVCAFEALAEAGVAKMGMPSSLAHRTAOTLLGTAK  
MLLHGOHPAQLRSDVCPGPTTIYGLHLEGGGLRAATMSAVERATCRKELSRK"

## BASE COUNT

ORIGIN

Query Match 84.3%; Score 1127.6; DB 9; Length 1178;  
Best Local Similarity 99.8%; Pred. No. 1.7e-192;  
Matches 1129; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

15 GTCGAGGCAACAAGATGCGAGCTGGGAGCGCTCTCCGCGGGCTGGGCTTGTGGGC 74  
|||||  
32 GTCGAGGCAACAAGATGCGAGCTGGGAGCGCTCTCCGCGGGCTGGGCTTGTGGGC 91  
|||||  
75 GCGGGCGCATGCGGGGGCCATCGCGCAGGGCCCTCATCAGAGCAGGAAAGTGAAGCT 134  
|||||  
92 GCGGGCGCATGCGGGGGCCATCGCGCAGGGCCCTCATCAGAGCAGGAAAGTGAAGCT 151  
|||||  
135 CAGCACATATGCGCCAGTGCACCAACAGACAGAGAACTATGTCACCTTCAAGCTCTGGGT 194  
|||||  
152 CAGCACATATGCGCCAGTGCACCAACAGACAGAGAACTATGTCACCTTCAAGCTCTGGGT 211  
|||||  
195 TGGCGGACACGACATCCCAACAGAGAGTGTGCAGAGCTGCCTGCTGCTCATCTTTGCC 254  
|||||  
212 TGGCGGACACGACATCCCAACAGAGAGTGTGCAGAGCTGCCTGCTGCTCATCTTTGCC 271  
|||||  
255 ACCAAGCCTCATGTGCGCAGCTGTCTGGCAGAGGTGGCTCTGTGTCTCACCAGTAA 314  
|||||  
272 ACCAAGCCTCATGTGCGCAGCTGTCTGGCAGAGGTGGCTCTGTGTCTCACCAGTAA 331  
|||||  
315 CACATCTGTGTGCTGCGGTGTCTCTGAGCACCCCTGGAGGAGTGTGTGCC 374  
|||||  
332 CACATCTGTGTGCTGCGGTGTCTCTGAGCACCCCTGGAGGAGTGTGTGCC 391  
|||||  
375 CCAACACACGGGTGTGCGGTGTCTGCCCAACCTGCCCTGTGTGTCTCAGGAGGGCC 434  
|||||  
392 CCAACACACGGGTGTGCGGTGTCTGCCCAACCTGCCCTGTGTGTCTCAGGAGGGCC 451  
|||||  
435 ATAGTGATGCGGGGGCCCGAGTGGGAGCAGCAGACCAAGCTCTGTCAGCATCTG 494  
|||||  
452 ATAGTGATGCGGGGGCCCGAGTGGGAGCAGCAGACCAAGCTCTGTCAGCATCTG 511  
|||||  
495 CTGAGGCGCTGTGCGGTGTGAGGAGGTGCCTGAAGCCCTACGTCGACATCCACACTGGC 554  
|||||  
512 CTGAGGCGCTGTGCGGTGTGAGGAGGTGCCTGAAGCCCTACGTCGACATCCACACTGGC 571  
|||||  
555 CTCAGTGGCAGTGGCGCTTTCGCTGTGTGTCATCTCCGAGGCCCTGGCTGAAGAGCC 614  
|||||  
572 CTCAGTGGCAGTGGCGCTTTCGCTGTGTGTCATCTCCGAGGCCCTGGCTGAAGAGCC 631  
|||||

|   |     |   |  |                |  |           |  |         |  |          |  |
|---|-----|---|--|----------------|--|-----------|--|---------|--|----------|--|
| BASE COUNT  |     | 142 a   |  | 248 c          |  | 264 g     |  | 133 t   |  | 7 others |  |
| ORIGIN  |     |   |  |                |  |           |  |         |  |          |  |
| Query Match   |     | 55.1%; Score 737.6; DB 6; Length 794;                               |  |                |  |           |  |         |  |          |  |
| Best Local Similarity   |     | 97.7%; Pred. No. 1.8e-122;  |  |                |  |           |  |         |  |          |  |
| Matches 765;  |     | Conservative 0;   |  | Mismatches 16; |  | Indels 2; |  | Gaps 2; |  |          |  |
| QY  | 15  | GTCGAGGCAACAAGATGGCAGCTGCGAGCGCGTCTCCGCGCGCGCTGGGCTTCGTGGGC 74      |  |                |  |           |  |         |  |          |  |
| Db  |     |   |  |                |  |           |  |         |  |          |  |
| QY  | 12  | GTNCGAGGCAACAAGATGGCAGCTGCGAGCCGINTCCGCGCGCTGGGCTTCGTGGGC 71        |  |                |  |           |  |         |  |          |  |
| Db  |     |   |  |                |  |           |  |         |  |          |  |
| QY  | 75  | GCGGGCCCATGCGGGGGCCATCGCGAGGGCCTCATCAGAGCAGGAAAGTGGAAAGCT 134       |  |                |  |           |  |         |  |          |  |
| Db  |     |   |  |                |  |           |  |         |  |          |  |
| QY  | 72  | GCGGGCCCATGCGGGGGGGCCATCGCGAGGGCCTCATCAGAGCAGGAAAGTGGAAAGCT 131     |  |                |  |           |  |         |  |          |  |
| Db  |     |   |  |                |  |           |  |         |  |          |  |
| QY  | 135 | CAGCACATACCTGGCCAGTGCCACCAACAGACAGAGAACCTATGTCACTTTCAGAGCTCTGGT 194 |  |                |  |           |  |         |  |          |  |
| Db  |     |   |  |                |  |           |  |         |  |          |  |
| QY  | 132 | CAGCACATACCTGGCCAGTGCCACCAACAGACAGAGAACCTATGTCACTTTCAGAGCTCTGGT 191 |  |                |  |           |  |         |  |          |  |
| Db  |     |   |  |                |  |           |  |         |  |          |  |
| QY  | 195 | TGCCGGACAGCGACTCCAACCCAGGAGGTGCTGCAGAGCTGCCGTGCTCATCTTTGCC 254      |  |                |  |           |  |         |  |          |  |
| Db  |     |   |  |                |  |           |  |         |  |          |  |
| QY  | 192 | TGCCGGACAGCGACTCCAACCCAGGAGGTGCTAGCAGAGCTGCCGTGCTCATCTTTGCC 251     |  |                |  |           |  |         |  |          |  |
| Db  |     |   |  |                |  |           |  |         |  |          |  |
| QY  | 255 | ACCAAGCCTCATGTGCTGCCAGCTGCTCTGSCAGAGTGCTCTGTGTTGGTCAACACACTGAA 314  |  |                |  |           |  |         |  |          |  |
| Db  |     |   |  |                |  |           |  |         |  |          |  |
| QY  | 252 | ACCAAGCCTCATGTGCTGCCAGCTGCTCTGSCAGAGTGCTCTGTGTTGGTCAACACACTGAA 311  |  |                |  |           |  |         |  |          |  |
| Db  |     |   |  |                |  |           |  |         |  |          |  |
| QY  | 315 | CACATCTTGTGTCCGTGGCTGTCTGTAGCACCCCTGGAGGAGCTGCTGCC 374              |  |                |  |           |  |         |  |          |  |
| Db  |     |   |  |                |  |           |  |         |  |          |  |
| QY  | 312 | CACATCTTGTGTCCGTGGCTGTCTGTAGCACCCCTGGAGGAGCTGCTGCC 371              |  |                |  |           |  |         |  |          |  |
| Db  |     |   |  |                |  |           |  |         |  |          |  |
| QY  | 375 | CCAAACACACGGGTGCTGCGGTCTTGCCCAACCTGCCCTGTGTGTCTCAGGAAGGGGCC 434     |  |                |  |           |  |         |  |          |  |
| Db  |     |   |  |                |  |           |  |         |  |          |  |
| QY  | 372 | CCAAACACACGGGTGCTGCGGTCTTGCCCAACCTGCCCTGTGTGTCTCAGGAAGGGGCC 431     |  |                |  |           |  |         |  |          |  |
| Db  |     |   |  |                |  |           |  |         |  |          |  |
| QY  | 435 | ATAGTGATGGCGGGGCCGCCACGCTGGGGAGCAGCAGACCAAGCTCC-TGCAGCATCT 493      |  |                |  |           |  |         |  |          |  |
| Db  |     |   |  |                |  |           |  |         |  |          |  |
| QY  | 432 | ATAGTGATGGCGGGGCCGCCACGCTGGGGAGCAGCAGACCAAGCTCCTTGCAGCATCT 491      |  |                |  |           |  |         |  |          |  |
| Db  |     |   |  |                |  |           |  |         |  |          |  |
| QY  | 494 | GCTGGAGCCTGTGGCGGTGTGAGGAGTGCCTGAAGCCTAGTCACATTCACACTGG 553         |  |                |  |           |  |         |  |          |  |
| Db  |     |   |  |                |  |           |  |         |  |          |  |
| QY  | 492 | GCTGGAGCCTGTGGCGGTGTGAGGAGTGCCTGAAGCCTAGTCACATTCACACTGG 551         |  |                |  |           |  |         |  |          |  |
| Db  |     |   |  |                |  |           |  |         |  |          |  |
| QY  | 554 | CCTCAGTGGCAGTGGCTTGCCCTTCGTGTGTGCATTCTCCGAGGCCCTGCTGAAGGAGC 613     |  |                |  |           |  |         |  |          |  |
| Db  |     |   |  |                |  |           |  |         |  |          |  |
| QY  | 552 | CCTCAGTGGCAGTGGCTTGCCCTTCGTGTGTGCATTCTCCGAGGCCCTGCTGAAGGAGC 611     |  |                |  |           |  |         |  |          |  |
| Db  |     |   |  |                |  |           |  |         |  |          |  |
| QY  | 614 | CGTCAAGATGGGATGCCACAGACGCTTGCCCCACCGCATCGCTGCCAGACCTGCTGGG 673      |  |                |  |           |  |         |  |          |  |
| Db  |     |   |  |                |  |           |  |         |  |          |  |
| QY  | 612 | CGTCAAGATGGGATGCCACAGANNCTGGCCCCACCGCATCGCTGCCAGACCTGCTGGG 671      |  |                |  |           |  |         |  |          |  |
| Db  |     |   |  |                |  |           |  |         |  |          |  |
| QY  | 674 | GACGGCCAAGATGCTGTCACAGAGGCCAACACCCAGCCAGCTGGCTCAGACGTGTG 733        |  |                |  |           |  |         |  |          |  |
| Db  |     |   |  |                |  |           |  |         |  |          |  |
| QY  | 672 | GACGGCCAAGATGCTGTCACAGAGGCCAACACCCAGCCAGCTGNGCTTCAGACGTGTG 731      |  |                |  |           |  |         |  |          |  |
| Db  |     |   |  |                |  |           |  |         |  |          |  |
| QY  | 734 | CA-CCCCGGGTGGCACCACCATCATATGGACTCCACGCCCTTGAGCAGGGCGGGCTGCGAG 792   |  |                |  |           |  |         |  |          |  |
| Db  |     |   |  |                |  |           |  |         |  |          |  |
| QY  | 732 | CACCCCGGTGGGACCAACCATCATATGGACTCCACGCCCTTGGAACAGGGCCGGCTGNA 791     |  |                |  |           |  |         |  |          |  |
| Db  |     |   |  |                |  |           |  |         |  |          |  |
| QY  | 793 | CAG 795   |  |                |  |           |  |         |  |          |  |
| Db  |     |   |  |                |  |           |  |         |  |          |  |
| QY  | 792 | CAG 794   |  |                |  |           |  |         |  |          |  |
| Db  |     |   |  |                |  |           |  |         |  |          |  |
| RESULT 8  |     |   |  |                |  |           |  |         |  |          |  |
| AC105210/c  |     |   |  |                |  |           |  |         |  |          |  |
| LOCUS AC105210 101777 bp DNA linear PRI 27-MAR-2003                         |     |   |  |                |  |           |  |         |  |          |  |
| DEFINITION Homo sapiens chromosome 8, clone RP11-118P12, complete sequence. |     |   |  |                |  |           |  |         |  |          |  |
| ACCESSION AC105210  |     |   |  |                |  |           |  |         |  |          |  |
| VERSION AC105210.10 GI:29294339   |     |   |  |                |  |           |  |         |  |          |  |
| KEYWORDS HTG.   |     |   |  |                |  |           |  |         |  |          |  |









\* 37554 38694: contig of 1141 bp in length  
\* 38595 38794: gap of 100 bp  
\* 38795 39930: contig of 1136 bp in length  
\* 39931 40030: gap of 100 bp  
\* 40031 41742: contig of 1712 bp in length  
\* 41743 41842: gap of 100 bp  
\* 41843 43678: contig of 1836 bp in length  
\* 43679 43778: gap of 100 bp  
\* 43779 45606: contig of 1828 bp in length  
\* 45607 47063: gap of 100 bp  
\* 47064 47863: contig of 2157 bp in length  
\* 47864 47963: gap of 100 bp  
\* 47964 50045: contig of 2082 bp in length  
\* 50046 50145: gap of 100 bp  
\* 50146 51025: contig of 880 bp in length  
\* 51026 51125: gap of 100 bp  
\* 51126 52693: contig of 1568 bp in length  
\* 52694 52793: gap of 100 bp  
\* 52794 55531: contig of 2738 bp in length  
\* 55532 55631: gap of 100 bp  
\* 55632 57290: contig of 1659 bp in length  
\* 57291 57390: gap of 100 bp  
\* 57391 58997: contig of 1607 bp in length  
\* 58998 59097: gap of 100 bp  
\* 59098 60543: contig of 1446 bp in length  
\* 60544 60643: gap of 100 bp  
\* 60644 62329: contig of 1686 bp in length  
\* 62330 62429: gap of 100 bp  
\* 62430 64069: contig of 1640 bp in length  
\* 64070 64169: gap of 100 bp  
\* 64170 66375: contig of 2206 bp in length  
\* 66376 66475: gap of 100 bp  
\* 66476 68972: contig of 2497 bp in length  
\* 68973 69072: gap of 100 bp  
\* 69073 70628: contig of 1556 bp in length  
\* 70629 70728: gap of 100 bp  
\* 70729 73435: contig of 2707 bp in length  
\* 73436 73535: gap of 100 bp  
\* 73536 75771: contig of 2236 bp in length  
\* 75772 75871: gap of 100 bp  
\* 75872 78176: contig of 2305 bp in length  
\* 78177 78276: gap of 100 bp  
\* 78277 80144: contig of 1868 bp in length  
\* 80145 80244: gap of 100 bp  
\* 80245 82363: contig of 2119 bp in length  
\* 82364 82463: gap of 100 bp  
\* 82464 84226: contig of 1763 bp in length  
\* 84227 84326: gap of 100 bp  
\* 84327 87431: contig of 3105 bp in length  
\* 87432 87531: gap of 100 bp  
\* 87532 90276: contig of 2745 bp in length  
\* 90277 90376: gap of 100 bp  
\* 90377 93023: contig of 2647 bp in length  
\* 93024 93123: gap of 100 bp  
\* 93124 95831: contig of 2708 bp in length  
\* 95832 95931: gap of 100 bp  
\* 95932 99011: contig of 3080 bp in length  
\* 99012 99111: gap of 100 bp  
\* 99112 101722: contig of 2611 bp in length  
\* 101723 101822: gap of 100 bp  
\* 101823 105034: contig of 3212 bp in length  
\* 105035 105134: gap of 100 bp  
\* 105135 108113: contig of 2979 bp in length  
\* 108114 108213: gap of 100 bp  
\* 108214 112103: contig of 3892 bp in length  
\* 112105 112205: gap of 100 bp  
\* 112206 115406: contig of 3201 bp in length  
\* 115407 115506: gap of 100 bp  
\* 115507 119714: contig of 4208 bp in length  
\* 119715 119814: gap of 100 bp  
\* 119815 125158: contig of 5344 bp in length  
\* 125159 125258: gap of 100 bp  
\* 125259 130255: contig of 4997 bp in length

\* 130256 130355: gap of 100 bp  
\* 130356 135663: contig of 5308 bp in length  
\* 135664 135763: gap of 100 bp  
\* 135764 141345: contig of 5582 bp in length  
\* 141346 141445: gap of 100 bp  
\* 141446 146381: contig of 4936 bp in length  
\* 146382 146481: gap of 100 bp  
\* 146482 155041: contig of 8560 bp in length  
\* 155042 155141: gap of 100 bp  
\* 155142 164370: contig of 9229 bp in length  
\* 164371 164470: gap of 100 bp  
\* 164471 164959: contig of 489 bp in length.  
FEATURES  
Location/Qualifiers  
Query Match 48.8%; Score 653.2; DB 2: Length 164959;  
Best Local Similarity 98.1%; Pred. No. 8.2e-108;  
Matches 661; Conservative 0; Mismatches 13; Indels 0; Gaps 0;  
QY 664 CCTCTGGGACGCCCAAGATGCTGTGCACGAGGGCCAAACACCCAGCCAGCTGCCT 723  
Db 141067 CTCCTCAGGGAGGCCCAAGATGCTGTGCACGAGGGCCAAACACCCAGCCAGCTGCCT 141008  
QY 724 CAGAGTGTGACCCCGGGTGGCCACCACCATCTATGACTCCAGCCCTGAGCAGGGCG 783  
Db 141007 CAGAGTGTGACCCCGGGTGGCCACCACCATCTATGACTCCAGCCCTGAGCAGGGCG 140948  
QY 784 GGCTCGAGCAGCCACCACCATGAGCGCGTGGAGGCTGCCACCTGCCGGGCCAAGAGCTCA 843  
Db 140947 GGCTCGAGCAGCCACCACCATGAGCGCGTGGAGGCTGCCACCTGCCGGGCCAAGAGCTCA 140888  
QY 844 CGAGAAAGTAGGCTGGGCTCTGGCCATCCTTTCTCCCTCTGTGCCCCCTGCTCTCCCTG 903  
Db 140887 CGAGAAAGTAGGCTGGGCTCTGGCCATCCTTTCTCCCTCTGTGCCCCCTGCTCTCCCTG 140828  
QY 904 TGTCCTTCCCTGAGGACTGGGGTCCCTCCCTCCTCGATGAGGGTCTCTACTGCTCC 963  
Db 140827 TGTCCTTCCCTGAGGACTGGGGTCCCTCCCTCCTCGATGAGGGTCTCTACTGCTCC 140768  
QY 964 TTCTCCCTTGACACAGGAAATGACAGGGGACAGGACTTGGAGGTTCCAGCAGGGGGGG 1023  
Db 140767 TTCTCCCTTGACACAGGAAATGACAGGGGACAGGACTTGGAGGTTCCAGCAGGGGGGG 140708  
QY 1024 AGCCCCGACAGTGGGGACACTTCTCCCTCCCTCCAGTGAAGAGACACCGTGGTGGTG 1083  
Db 140707 AGCCCCGACAGTGGGGACACTTCTCCCTCCCTCCAGTGAAGAGACACCGTGGTGGTG 140648  
QY 1084 CTCTGCCCCCTGCTGAGTGAGCCACCTTGTGTCAACATTGGTTCTGAGGGGGCCCAAGA 1143  
Db 140647 CTCTGCCCCCTGCTGAGTGAGCCACCTTGTGTCAACATTGGTTCTGAGGGGGCCCAAGA 140588  
QY 1144 GATGGCTCTTGGTCAATTTGGCCGATGGTTGGGAGTGGTTGAGGGCATGAACAGAAC 1203  
Db 140587 GATGGCTCTTGGTCAATTTGGCCGATGGTTGGGAGTGGTTGAGGGCATGAACAGAAC 140528  
QY 1204 TTACGGTAACAGGCACGGCTGGCCCAATGCCTGGTGTGAGCTGAGCTTGCTTTGGCT 1263  
Db 140527 TTACGGTAACAGGCACGGCTGGCCCAATGCCTGGTGTGAGCTGAGCTTGCTTTGGCT 140468  
QY 1264 TTCCAAAGTGGGCTGCTGAGCTACAGCCAGGGCGGCTGCCTCATCTACCTCTAGGGGGC 1323  
Db 140467 TTCCAAAGTGGGCTGCTGAGCTACAGCCAGGGCGGCTGCCTCATCTACCTCTAGGGGGC 140408  
QY 1324 ACGAGCATATGGG 1337  
Db 140407 ACGAGCATATGGG 140394  
RESULT 10  
AC067930/c 174906 bp DNA linear PRI 24-JUL-2002  
LOCUS Homo sapiens chromosome 8, clone RP11-661A12, complete sequence.  
DEFINITION AC067930  
ACCESSION AC067930  
VERSION AC067930.7 GI:21954034

## KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 174906)

Birren,B., Nusbaum,C. and Lander,E.

Homo sapiens chromosome 8, clone RP11-661A12

2 (bases 1 to 174906)

Birren,B., Linton,L., Nusbaum,C., Lander,E., Abrahams,H., Allen,N.,

Anderson,S., Baldoni,J., Barna,N., Bastien,V., Beda,F.,

Boguslavsky,L., Boukhalter,B., Brown,A., Burkett,G.,

Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,

Collamore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S.,

Dodge,S., Domino,M., Doyle,M., Ferreira,P., Fitzhugh,W., Gage,D.,

Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,

Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,

Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,

Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczy,J.,

Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,

McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,

Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,

Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,

O'Neil,D., Oliver,J., Pollara,V., Raymond,C., Retta,R., Rogov,P.,

Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,

Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,

Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,

Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,

Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,

Young,G., Zainoun,J., Zimmer,A. and Zody,M.

Direct Submission

Submitted (27-APR-2000) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

3 (bases 1 to 174906)

Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,

Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L.,

Boukhalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,

Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collamore,A.,

Cooke,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,

Faro,S., Ferreira,P., Fitzgerald,M., Fitzhugh,W., Gage,D.,

Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L.,

Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I.,

Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K.,

Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Lindblad-Toh,K.,

Liu,G., Maclean,C., Macdonald,P., Major,J., Marquis,N.,

Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Meldrim,J.,

Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,

Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P.,

O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,

Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C.,

Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S.,

Schupback,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N.,

Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S.,

Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J.,

Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,

Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission

Submitted (21-JUN-2002) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

4 (bases 1 to 174906)

Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,

Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhalter,B.,

Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collamore,A.,

Cooke,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,

Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J.,

Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,

Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,

Karatas,A., Kells,C., Macdonald,P., Major,J., Matthews,C.,

Liu,G., Maclean,C., Macdonald,P., Meneus,L., Mihova,T., Mlenga,V.,

McCarthy,M., Meldrim,J., Meneus,L., Mlenga,V.,

Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,

O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,

## TITLE

JOURNAL

## COMMENT

Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission

Submitted (24-JUL-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

On Jul 24, 2002 this sequence version replaced gi:21490265.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997) <http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)

----- Project Information

Center project name: L10100

Center clone name: 661\_A\_12

-----

## FEATURES

Source

Location/Qualifiers

1..174906

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

/chromosome="8"

/map="8"

/clone="RP11-661A12"

/clone\_lib="RPCI-11 Human Male BAC"

/complement(351..430)

/rpt\_family="MIR"

977..1108

/rpt\_family="L1MC4a"

/complement(1149..1449)

/rpt\_family="AluSq"

3675..3781

/rpt\_family="MIR"

/complement(16671..16974)

/rpt\_family="AluSq"

20174..20412

/rpt\_family="L2"

20487..20585

/rpt\_family="L2"

/complement(20694..20702)

/note="<30 qual SNGL region"

23655..23772

/rpt\_family="L1ME2"

24264..24383

/rpt\_family="L1ME2"

/complement(24446..24779)

/rpt\_family="MLTib"

26300

/note="probably A; possibly G"

26681..26825

/rpt\_family="L1ME2"

28798..29091

/rpt\_family="AluSq"

32096..32118

/rpt\_family="AT\_rich"

32990..33157

/rpt\_family="L1M2"

33598..33749

/rpt\_family="L1ME2"

/complement(33750..34064)

/rpt\_family="AluJo"

34065..34129

/rpt\_family="L1ME2"

34132..34706

/rpt\_family="L1MC4a"

34991..35084

/rpt\_family="L1MD3"

## TITLE

JOURNAL

## REFERENCE

AUTHORS

## TITLE

JOURNAL

## REFERENCE

AUTHORS

|   |  |  |        |   |        |
|---|--|--|--------|---|--------|
| repeat_region   | 35191...35325                            | /rpt_family="LMC4"   | 143815 | TGTCCTTCCCTCGAGACTGCGGCTCCCTCCCTCCTCGCATGAGGGTCTCTACTGCTCC  | 143756 |
| repeat_region   | 36901...37069                            | /rpt_family="LMC4"   | 964    | TTCTCCCTTGCACAGGGAATGCAGGGGCGAGACTTGGGAGGTTCCAGCAGCGGGGG    | 1023   |
| repeat_region   | 37077...37162                            | /rpt_family="LMC4"   | 143755 | TTCTCCCTTGCACAGGGAATGCAGGGGCGAGACTTGGGAGGTTCCAGCAGCGGGGG    | 143696 |
| repeat_region   | 38672...38844                            | /rpt_family="MSTC"   | 1024   | AGCCCGGACACAGTGGGACACTCCTCCCTCCCGAGTGCAGAGAGCCCGTGGTGGTGG   | 1083   |
| unsure  | 38859...38884                            | /rpt_family="MSTC"   | 143695 | AGCCCGGACACAGTGGGACACTCCTCCCTCCCGAGTGCAGAGAGCCCGTGGTGGTGG   | 143636 |
| repeat_region   | 38880...38983                            | /rpt_family="MSTC"   | 1084   | CTCTGCCCTTGGCTGAGCCACCTTTGCTGCAACATTGGTTCTGAGGGGCCCAAGA     | 1143   |
| unsure  | 38996...39093                            | /rpt_family="MSTC"   | 143635 | CTCTGCCCTTGGCTGAGCCACCTTTGCTGCAACATTGGTTCTGAGGGGCCCAAGA     | 143576 |
| unsure  | 39033...39092                            | /rpt_family="MSTC"   | 1144   | GATGCGCTCTTGGTCATTTGCCCGCATGTTGGGCGAGTGGTTGAGGCCATGAACAGAAC | 1203   |
| repeat_region   | 39281...39281                            | /rpt_family="MSTC"   | 143575 | GATGCGCTCTTGGTCATTTGCCCGCATGTTGGGCGAGTGGTTGAGGCCATGAACAGAAC | 143516 |
| repeat_region   | 39282...39334                            | /rpt_family="MSTC"   | 1204   | TTACGGTAACAGGACCGCTGCGCCAAATGCTGCTGAGCTGGAGCTTGCCCTTTGGCT   | 1263   |
| repeat_region   | 39366...39668                            | /rpt_family="AluJb"  | 143515 | TTACGGTAACAGGACCGCTGCGCCAAATGCTGCTGAGCTGGAGCTTGCCCTTTGGCT   | 143456 |
| repeat_region   | 40044...40423                            | /rpt_family="MLT1A2"   | 1264   | TTCCAAGTGGGCTGCTGCAGCTACAGCCAGCGCGGCTGCCTCATCTCAGCTTAGGGGGC | 1323   |
| repeat_region   | 41752...41933                            | /rpt_family="MLT1A2"   | 143455 | TTCCAAGTGGGCTGCTGCAGCTACAGCCAGCGCGGCTGCCTCATCTCAGCTTAGGGGGC | 143396 |
| repeat_region   | 42478...42510                            | /rpt_family="G-rich"   | 1324   | ACGAGCATATGGG   | 1337   |
| unsure  | 43664...43676                            | /rpt_family="G-rich"   | 143395 | ACGAGCATATGGG   | 143382 |
| unsure  | 43674...43674                            | /note="<30 qual SNGL region"                                   |        |   |        |
| unsure  | 43674...43674                            | /note="probably ACTGAG"  |        |   |        |
| unsure  | 43674...43674                            | /note="probably ACTGGTGGC"                                     |        |   |        |
| repeat_region   | 46481...46483                            | /note="probably CCC"   |        |   |        |
| repeat_region   | 54691...54749                            | /rpt_family="(CA)n"  |        |   |        |
| repeat_region   | 56353...56388                            | /rpt_family="(CA)n"  |        |   |        |
| repeat_region   | 60486...60780                            | /rpt_family="MER5B"  |        |   |        |
| repeat_region   | 60802...61091                            | /rpt_family="AluSx"  |        |   |        |
| repeat_region   | 69122...69423                            | /rpt_family="AluJb"  |        |   |        |
| repeat_region   | 71273...71458                            | /rpt_family="AluJb"  |        |   |        |
| repeat_region   | 71493...71733                            | /rpt_family="L2"   |        |   |        |
| repeat_region   | 71849...72140                            | /rpt_family="AluSx"  |        |   |        |
| Query Match   | 48.8%; Score 653.2; DB 9; Length 174906; |  |        |   |        |
| Best Local Similarity   | 98.1%; Pred. No. 8.1e-108;               |  |        |   |        |
| Matches 661; Conservative 0; Mismatches 13; Indels 0; Gaps 0; |  |  |        |   |        |
| QY  | 664                                      | CCCTGCTGGGACGCGCCAAAGATGCTGTGCACGAGGGCCCAACACCCAGCCCGAGCTGCGCT | 723    |   |        |
| DB  | 144055                                   | CTCCTCAGGGACGCGCAAGATGCTGTGCACGAGGGCCCAACACCCAGCCCGAGCTGCGCT   | 143996 |   |        |
| QY  | 724                                      | CAGAGGTGTGACCCCGGGTGGACACCATCTATGACTTCACGCGCTGGAGGGGGG         | 783    |   |        |
| DB  | 143995                                   | CAGAGGTGTGACCCCGGGTGGACACCATCTATGACTTCACGCGCTGGAGGGGGG         | 143936 |   |        |
| QY  | 784                                      | GCCTGGAGCAGCCACCATGAGCGCGCTGGAGGCTGCCACCTGCCGGGCCAAGAGCTCA     | 843    |   |        |
| DB  | 143935                                   | GCCTGGAGCAGCCACCATGAGCGCGCTGGAGGCTGCCACCTGCCGGGCCAAGAGCTCA     | 143876 |   |        |
| QY  | 844                                      | GCAGAAAGTAGGCTGGGCTCTGGGCGATCCCTTTCTGCTCTGTGCGCCCTGCTCCCTG     | 903    |   |        |
| DB  | 143875                                   | GCAGAAAGTAGGCTGGGCTCTGGGCGATCCCTTTCTGCTCTGTGCGCCCTGCTCCCTG     | 143816 |   |        |
| QY  | 904                                      | TGTCCTTCCCTGAGGACTGGGCTCCCTCCCTTCCTGATGAGGGTCTCTACTGCTCC       | 963    |   |        |

|    |       |   |       |
|----|-------|---|-------|
| QY | 664   | CCCTGCTGGGAGCGGCCAAGATGCTGTGCACAGAGGGCCAAACCCAGAGCCCAAGTCGCCT   | 723   |
| Db | 48388 | CTCTCAGAGGGAGGGCCAAAGATGCTGCTGCACAGAGGGCCAAACCCAGAGCCCAAGTCGCCT | 48447 |
| QY | 724   | CAGAGGTGTGCACCCCGGTGGCACACCACTATGAGCTCCACGCCCTGGAGCAGGGCG       | 783   |
| Db | 48448 | CAGAGGTGTGCACCCCGGTGGCACCACTATGAGCTCCACGCCCTGGAGCAGGGCT         | 48507 |
| QY | 784   | GGCTGCGAGCAGCCACCATGAGCGCGCTGGAGGCTGCCACCTCCGGGGCCAAAGAGCTCA    | 843   |

|       |    |   |       |
|-------|----|---|-------|
| 48508 | Db | GNCTGCCAGCAG-CACCATGAGCGCGTGGAGG-TGCCACCTCCGGGCGAAGGAGCTCA      | 48565 |
| 844   | QY | GCAGAAAGTAGCTGGGCTCTGCCCATCTTTTCCTGCTCTGTGTGCCCTCGCTCTCCCTG     | 903   |
| 48566 | Db | GCAGAAAGTAGCTGGGCTCTGG-CATCTTTCTGCTCTGTGTGCCCTCGCTCTCCCTG       | 48624 |
| 904   | QY | TGTCCTCTCCCTCGAGCACTGCGCTCCCTCCCTCTCTGCATGATGAGGCTCTCTACTGTCTCC | 963   |
| 48625 | Db | TGTCTCTCTCCCTCGAGCACTGCGGCTCCCTCCCTCTCTGCATGAGGCTCTCTACTGTCTCC  | 48684 |
| 964   | QY | TTCTCTCCCTTGCCACAGGGAAATGCAGGGGGCAGGACTTGGAGGTTCCAGCAGCGGGGG    | 1023  |
| 48685 | Db | TTCTCTCCCTTGCCACAGGGAAATGCAGGGGGCAGGACTTGGAGGTTCCAGCAGCGGGGG    | 48744 |
| 1024  | QY | AGCCCCGACCAGTGGGGACACTCTCTCCCTCCAGTGAGCAGAGAAGCACCGTGGTGGTG     | 1083  |
| 48745 | Db | AG-CCGACCAAGTGGGGACACTCTCTCCCTCCAGTGAGCAGAGAAGCACCGTGGTGGTG     | 48803 |
| 1084  | QY | CTCTGCCCCCTTGCTGCAGTGAGCCCACTTGCCTCAACATTTGTTGTGAGGGGCCCAAGA    | 1143  |
| 48804 | Db | CTCTGCCCCCTTGCTGCAGTGAGCCCACTTGCCTCAACATTTGTTGTGAGGGGCCCAAGA    | 48863 |
| 1144  | QY | GATGGGCTCTTGCTCATTTTGGCCCGCATGGTTGGGCAGTTGGTTGAGGCCCATGAACAGAAC | 1203  |

|    |       |  |       |
|----|-------|--|-------|
| QY | 1144  | GATGGCGCTCTGGTGCATTTGCCCGCATGGTGGGCACTGGTGGAGCCATGAACAGAAC   | 48923 |
| Db | 48864 | GATGGCGCTCTGGTGCATTTGCCCGCATGGTGGGCACTGGTGGAGCCATGAACAGAAC   |       |
| QY | 1204  | TTACGGTAAACAGGCACGGCTGGCCCAATGCCTGGTCTGGAGCTGGAGCTTGCTTTGGCT | 1263  |
| Db | 48924 | TTACGGTAAACAGGCACGGCTGGCCCAATGCCTGGTCTGGAGCTGGAGCTTGCTTTGGCT | 48983 |

|            |   |  |       |
|------------|---|--|-------|
| Qy         | 1264  | TTCACAGTGGGCTCGTGCACTACAGCCAGGCCGGCTGCCTCATCTCAGCTCTAGGGGC | 1323  |
| Db         | 48984   | TTTCCAGTGGCTCCGTGCACTACAGCCAGGCCGGCTGCCTCATCTCAGCTCTAGGGGC | 49043 |
| Qy         | 1324  | ACGAGCATATGGG  | 1337  |
| Db         | 49044   | ACGAGCCATATGGG   | 49057 |
| RESULT     | 12  |  |       |
| LOCUS      | BC026536  |  |       |
| LOCUS      | Mus musculus, RIKEN cDNA I110058B13 gene, clone MGC:35759   | 1276 bp mRNA linear ROD 20-SEP-2002                        |       |
| DEFINITION | IMAGE:4987934, mRNA, complete cds.                          |  |       |
| ACCESSION  | BC026536  |  |       |
| VERSION    | BC026536.1  | GI:20071633  |       |
| KEYWORDS   | MGC.  |  |       |
| SOURCE     | Mus musculus (house mouse)                                  |  |       |
| ORGANISM   | Mus musculus  |  |       |
| REFERENCE  | Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;     |  |       |
| AUTHORS    | Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.   |  |       |
| TITLE      | 1 (bases 1 to 1276)<br>Straussberg, R.<br>Direct Submission |  |       |

REMARK  
COMMENT

Submitted (02-APR-2002) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA  
NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [cgabbs-r@mail.nih.gov](mailto:cgabbs-r@mail.nih.gov)  
Tissue Procurement: Jeffrey E. Green, M.D.

COMMENT  
Contact: MGC help desk  
Email: [cgabs-r@mail.nih.gov](mailto:cgabs-r@mail.nih.gov)  
Tissue Procurement: Jeffrey E. Green, M.D.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Sequencing Group at the Stanford Human Genome  
Center, Stanford University School of Medicine, Stanford, CA 94305  
Web site: <http://www.shgc.stanford.edu>  
Contact: (Dickson, Mark) mcdpaxil.stanford.edu  
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,  
R. M.

Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAC Place: 58 Row: d Column: 3  
This clone was selected for full length sequencing because it  
passed the following selection criteria: matched mRNA gi: 13384799.

FEATURES  
source  
1..1276  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="FVB/N"  
/db\_xref="taxon:10090"  
/clone="MGC:35759 IMAGE:4987954"  
/tissue\_type="Colon, normal. 5 month old male mouse."  
/clone\_lib="NCI CGAP\_Co24"  
/lab\_host="DH10B"  
/note="Vector: pCMV-SPORT6"  
16..840  
/codon\_start=1  
/product="RIKEN cDNA 1110058B13 gene"  
/protein\_id="AAH26536.1"  
/db\_xref="GI:20071634"  
/db\_xref="LocusID:66194"  
/translation="MAATMSEPRRVGVFVAGRMATIAIQLQKVEAKOVLASAPT  
DNMLFRALGQTHSNHEVLQNPVIFATKQVLTVAEVAPIVTHIIVSA  
AGLSTMGVLPNPNVLRVPCVQVQCAMVARGHAGNDLLEQLLEACG  
OCLEVPESVDIHTGLSGSVAFVCTFEALAEAGAKMGPGLHRAQLLGTAK  
MLQEQEKHPAQLRTDVLTPAGTTIHLHALERGGFRAATMSAVEAATCRAKELSKK"  
288 a 357 c 360 g 271 t

CDS

BASE COUNT  
ORIGIN  
Query Match 42.4%; Score 567.6; DB 10; Length 1276;  
Best Local Similarity 79.8%; Pred. No. 5.3e-92;  
Matches 669; Conservative 0; Mismatches 169; Indels 0; Gaps 0;  
QY 24 AACAAAGTGGACGTGGGAGCGTCTCCGCGCGCGTGGTGGTGGTGGGCGGCGCG 83  
DB 10 AGCAGGATGGCGGCGACAATGTCTGAACCTCGGCGTGGTGGTGGTGGTGGTGG 69  
QY 84 ATGCGGGGGCCATCGCGAGCGCTCATCAGACGAGGAAAGTGAAGCTCAGCACATA 143  
DB 70 ATGCGGAGGCGATGCGCGGCGCTCATCAGACGAGGAAAGTGAAGCTCAGCACATA 129  
QY 144 CTGCGCAGTGCCACACAGACAGAACCTATGTCACTTTCAAGCTCTGGGTTGCCGACC 203  
DB 130 CTGCGCAGTGCCACACAGGACACACACCTCTGCCACTTCAGGCTCTGGTGGCAGACT 189  
QY 204 AGCAGCTCCACACAGGAGGTGCTGACAGAGTGGCTGTCTATCTTTGCCACCAAGCT 263  
DB 190 ACTCACTCCACACAGGAGGTGCTGCAAAACTGCGCACTTGTCTATCTTTGCCACCAACC 249  
QY 264 CATGTGTGCGAGCTGCTCGGAGAGTGGCTGTGTCACTTCACTTCACTTCACTTCACT 323  
DB 250 CAAGTCTTGCACACTGCTTGGGGAAGTGGGCGGCAATGATGATGATGATGATGATG 309  
QY 324 GTGTGCGGCTGCTGGGCTGTCTGTGAGCAGCCTGGAGGAGTGTGTGCGCCCAACACA 383  
DB 310 GTATCTGTGGTGTGGGATCTCTCTGAGCACAATGGGGGCTGTATCCCGGCAACACA 369  
QY 384 CGGTGTGCTGGGGTGTGGCCCAACCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 443  
DB 370 CGAGTATGCGAGTCTCTCCCAATCTACCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 429  
QY 444 GCGGGGGCGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 503  
DB 430 GCGCGGGGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 489

QY 504 TGTGGCGCGTGTGAGGAGTCCCTGAAGCTACGTGACATCCACATGCTGCTCAGTGGC 563  
DB 490 TGTGGCGCATGATAGAGTTCCCGAGTCTACGTATATCCACACCGTCTCAGTGGC 549  
QY 564 AGTGGCGTGGCTTGT 623  
DB 550 AGTGGTGTGGCTTGT 609  
QY 624 GGCATGCCAGCAGCGCTGGCCACCGCATCCAGAGCCCTGCTGGGAGCGGCCAAG 683  
DB 610 GGCATGCCAGTGGCTGGCCACCGCATTCGTCTCAGACCCCTGCTGGGAGCAGCCAA 669  
QY 684 ATGCTGTGTGACAGGCGGCAACACCCAGCCAGTGGCTCAGACGTGTGACACCCCGGT 743  
DB 670 ATGCTGTGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 729  
QY 744 GGCATGCCATCTATGGACTCCACGCTTGGAGCGGCGGCTGCGAGCAGCCACCATG 803  
DB 730 GGCATGCCATCTATGGACTCCACGCTTGGAGCGGCGGCTGCGAGCAGCCACCATG 789  
QY 804 AGCGCGTGGAGGCTGCCACCTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 861  
DB 790 AGTGGCGTGAAGCAGCTACCTGCGCGGCTAAGGAGCTCAGCAAGAGTGAAGGCGG 847  
RESULT 13  
BD148256 563 bp DNA linear PAT 17-JAN-2003  
LOCUS  
Primer for synthesizing full-length cDNA and use thereof.  
DEFINITION  
ACCESSION  
BD148256  
VERSION  
BD148256.1 GI:27854014  
KEYWORDS  
JP 2002191363-A/3099.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
1 (bases 1 to 563)  
Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,  
Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.  
Primer for synthesizing full-length cDNA and use thereof  
Patent: JP 2002191363-A 3099 09-JUL-2002;  
JOURNAL  
HELEX RESEARCH INSTITUTE  
COMMENT  
OS Homo sapiens (human)  
PN JP 2002191363-A/3099  
PD 09-JUL-2002  
PF 28-JUL-2000 JP 2000280990  
PI TOSHIO OTA,TAKAO ISOGAI,TETSUO NISHIKAWA,KOJI HAYASHI,KAORU  
PI SAITO,  
PI JUNICHI YAMAMOTO,SHIZUKO ISHII,TOMOYASU SUGIYAMA,AI WAKAMATSU,  
PI KEIICHI NAGAI,TETSUJI OTSUKI  
PC  
C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/ PC  
10,  
C12P21/02,C12P21/68/C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC  
Primer for synthesizing full-length cDNA and use thereof FH Key  
FEATURES  
source  
1..563  
Location/Qualifiers  
FT source  
1..563  
/organism="Homo sapiens (human)"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
BASE COUNT  
ORIGIN  
99 a 171 c 181 g 108 t 4 others  
Query Match 36.4%; Score 486.6; DB 6; Length 563;  
Best Local Similarity 95.9%; Pred. No. 2.1e-77;  
Matches 540; Conservative 0; Mismatches 18; Indels 5; Gaps 4;  
QY 27 AAGATGGCAGTGC-GGAGCGCTCTCCGCGCGCGTGGCTTGTGTGTGTGTGTGTGTGT 85  
|||||

Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,  
Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,  
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,  
Lorensuhewa,L., Louisseg,H., Lozado,R.J., Lu,X., Ma,J.,  
Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,  
Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,  
Mawhinney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,  
Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,  
Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,  
Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,  
Nwaekelehah,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,  
Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,  
Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.,  
Puzo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,  
Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,  
Rives,C., Rodkey,T., Rojdas,A., Rose,M., Rose,R., Ruiz,S.J.,  
Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,  
Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajls,D.,  
Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,  
Steinle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,  
Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,  
Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B.,  
Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,  
Williams,G., Willson,R., Wluczky,R., Wooden,H., Worley,K.,  
Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,D., Yoon,V.,  
Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von  
Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,  
Weinstock,G. and Gibbs,R.A.  
Direct Submission  
Unpublished  
2 (bases 1 to 197243)  
2  
Worley,K.C.  
Direct Submission  
Submitted (09-MAY-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 197243)  
Rat Genome Sequencing Consortium.  
Direct Submission  
Submitted (19-NOV-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
On Nov 19, 2002 this sequence version replaced gi:2856280.  
The sequence in this assembly is a combination of BAC based reads  
and whole genome shotgun sequencing reads assembled using Atlas  
(<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described  
in the feature table below represents a scaffold in the Atlas  
assembly (a 'contig-scaffold'). Within each contig-scaffold,  
individual sequence contigs are ordered and oriented, and separated  
by sized gaps filled with Ns to the estimated size. The sequence  
may extend beyond the ends of the clone and there may be sequence  
contigs within a contig-scaffold that consist entirely of whole  
genome shotgun sequence reads. Both end sequences and whole genome  
shotgun sequence only contigs will be indicated in the feature  
table.  
----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: hgsc-help@bcm.tmc.edu  
----- Project Information  
Center project name: GWMH  
Center clone name: CH230-300L19  
----- Summary Statistics  
Assembly program: Phrap; version 0.990329  
Consensus quality: 178119 bases at least Q40  
Consensus quality: 179916 bases at least Q30  
Consensus quality: 160768 bases at least Q20  
Estimated insert size: 180558; sum-of-contigs estimation  
Quality coverage: 8x in Q20 bases; sum-of-contigs estimation  
-----  
\* NOTE: Estimated insert size may differ from sequence length  
\* (see [http://www.hgsc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)).



|                       |  |   |   |          |
|-----------------------|--|---|---|----------|
| BASE COUNT            | 364 a                                  | 582 c   | 541 g   | 361 t    |
| ORIGIN                |  |   |   |          |
| Query Match           | 16.9%; Score 226.4; DB 9; Length 1848; |   |   |          |
| Best Local Similarity | 58.1%; Pred. No. 8e-31;                |   |   |          |
| Matches               | 467; Conservative                      | 0; Mismatches   | 316; Indels   | 21; Gaps |
|                       | 3;                                     |   |   |          |
| Qy                    | 56                                     | GC  | CGTGGGCTTCGTGGGCGCGCCGATGCGGGGGCCATGCGCGAGGGGCTCATCAG | 115      |
| Db                    | 52                                     | GAGCGTGGGCTTCATCGGCGCTGCGCAGCTGGCTTTTGGCTGGCCAAAGGCTTCACAGC   | 111   |          |
| Qy                    | 116                                    | AGCAGGAAGAGTGGAGCTCAGCACATAGTGGCCAGTGACCAACACAGAGAACTATG      | 175   |          |
| Db                    | 112                                    | AGCAGGCGTCTTGCTGCGCCACAAAGATAATGGCTAGCTCCC---CAGACATGACCTGGC  | 168   |          |
| Qy                    | 176                                    | TCACCTTCAAGCTC-----TGGGTTGCGGACACAGCACTCCAAACCAGGAGGTGCT      | 226   |          |
| Db                    | 169                                    | CACAGTTTCGTCTCTCAGGAGATGGGGTGAAGTTGACACCCACAAAGAGACGCT        | 228   |          |
| Qy                    | 227                                    | GCAGAGTGCCTGCTCGTCACTTTGCCAACAGCCTCATGTGCTGCCAGCTGTCTCTGGC    | 286   |          |
| Db                    | 229                                    | GCAGCACAGTGATGTCTCTTCCTGCTGTGAAGCCACACATCATCCCTTCATCCTGGA     | 288   |          |
| Qy                    | 287                                    | AGAGTGGCTTCTGTGGTGCACACTGAACACATCTTGGTGTGCGTGGCTGTGCGGGTGTG   | 346   |          |
| Db                    | 289                                    | TGAATAGGCGCGCACATTGAGGACACACATTTGTGTCTCTGCGCGCGCGGCTCAC       | 348   |          |
| Qy                    | 347                                    | TCGTGACACCTCGTGGAGGCTGTG-----CCCCCAACACACAGGGTGTGCGGGT        | 397   |          |
| Db                    | 349                                    | CATCAGTCCATTGAGAGAGAGTGTACGGTTTCGGCCAGCCCCCAGGGTTCATCGCTG     | 408   |          |
| Qy                    | 398                                    | CTTGCCCAACTCCCTGTGTGTCCAGGAAGGGCCATAGTATGGCGCGGGGCGCCA        | 457   |          |
| Db                    | 409                                    | CATGACCAACACTCCAGTCTGTGTGCGGAGGGGCGCACCGTGTATGCCACAGGCACGCA   | 468   |          |
| Qy                    | 458                                    | CTGCGGAGCAGGAGACACCACTCTCTGAGCATCTGTGGAGGCTGTGGGGGTGTGA       | 517   |          |
| Db                    | 469                                    | CGCCCCAGGTGGAGGAGGGAGGCTCATGGACAGCTGCTGAGCAGCTGGGGCTTCTGCAC   | 528   |          |
| Qy                    | 518                                    | GGAGTGCCTGAAGCCTTACGTCACATCCACACTGGCCTCAGTGGCAGTGGGCTGGCCTT   | 577   |          |
| Db                    | 529                                    | GGAGTGGAGAGAGACCTGATTGATGCCGTCACGGGGCTCAGTGGCAGCGGGCCCCCTA    | 588   |          |
| Qy                    | 578                                    | CGTGTGTGCATTTCCGAGGCCCTGTGCTGAAGAGCCGTCAGATGGGATGCCACAGCAG    | 637   |          |
| Db                    | 589                                    | CGCATTCACAGCCTGGATGCCCTGTGCTGATGGGGCGTGAAGATGGGACTTCCAAGGG    | 648   |          |
| Qy                    | 638                                    | CCTGGCCCCACCGCATCGCTGCCAGACCTGTCTGGGGACGGCCCAAGATGCTGCTGCACGA | 697   |          |
| Db                    | 649                                    | CCTGGCAGTCCGGCTCGGGCCCCAGGCCCTCTCTGGGGGTGCCAAGATGCTGCTGCATC   | 708   |          |
| Qy                    | 698                                    | GGGCCAACCCAGCCCGCTGGCTCAGACGTGTGTCGCCGGGTGGCACCACCATCTA       | 757   |          |
| Db                    | 709                                    | AGACACAGCACCCAGGCCAGCTCAAGACACACGTCAGCTCTCTCTGGTGGGGCCACCATCA | 768   |          |
| Qy                    | 758                                    | TGACTCCAGCCCTGGAGAGGGCGGGTGGAGAGCCACCATGAGCGCGCTGGAGGC        | 817   |          |
| Db                    | 769                                    | TGCTTGCATGTCTGGAGAGTGGGGCTTCCGCTCCCTGCTCATCAACGCTGTGGAGGC     | 828   |          |
| Qy                    | 818                                    | TGCCACCTGCGGGCCAGGAGCT  | 841   |          |
| Db                    | 829                                    | CTCTGTCATCCGACACAGGAGCT                                       | 852   |          |

Search completed: August 8, 2003, 16:36:16  
Job time : 4875 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 8, 2003, 14:36:31 ; Search time 338 Seconds  
(without alignments)  
8166.553 Million cell updates/sec

Title: US-09-806-536A-29  
Perfect score: 1338  
Sequence: 1 ggtgagcagctgtctgcga.....ggggcagcagcatatggggt 1338

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1439767 seqs, 1031500376 residues

Total number of hits satisfying chosen parameters: 2879534

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published\_Applications\_NA.:

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*
- 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*
- 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq.\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq.\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq.\*
- 11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq.\*
- 12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*
- 13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq.\*
- 14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq.\*
- 15: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*
- 16: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*
- 17: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score  | Query Match | Length  | ID | Description         |
|------------|--------|-------------|---------|----|---------------------|
| 1          | 1307   | 97.7        | 2331    | 14 | US-10-161-418A-2    |
| 2          | 1127.8 | 84.3        | 1178    | 14 | US-10-161-418A-7    |
| 3          | 857    | 64.1        | 860     | 14 | US-10-161-418A-6    |
| 4          | 226.4  | 16.9        | 1848    | 14 | US-10-161-418A-3    |
| 5          | 224.8  | 16.8        | 1792    | 10 | US-09-880-107-2400  |
| 6          | 224.8  | 16.8        | 1792    | 14 | US-10-161-418A-1    |
| 7          | 214.4  | 16.0        | 1715    | 14 | US-10-161-418A-5    |
| 8          | 211.2  | 15.8        | 1028    | 14 | US-10-161-418A-4    |
| 9          | 203.4  | 15.2        | 1478    | 10 | US-09-925-300-278   |
| 10         | 201.8  | 15.1        | 1708    | 14 | US-10-161-418A-8    |
| 11         | 201.8  | 15.1        | 1742    | 9  | US-09-912-717-2     |
| 12         | 167    | 12.5        | 999     | 14 | US-10-161-418A-9    |
| 13         | 156.4  | 11.7        | 807     | 14 | US-10-156-761-4708  |
| 14         | 156.4  | 11.7        | 9025608 | 14 | US-10-156-761-1     |
| 15         | 129    | 9.6         | 454     | 11 | US-09-918-995-11309 |
| 16         | 118.6  | 8.9         | 371     | 11 | US-09-918-995-181   |

|    |       |     |         |    |                      |                   |
|----|-------|-----|---------|----|----------------------|-------------------|
| 17 | 115.6 | 8.6 | 482     | 11 | US-09-918-995-31035  | Sequence 31035, A |
| 18 | 84.4  | 6.3 | 810     | 10 | US-09-738-626-460    | Sequence 460, App |
| 19 | 84.4  | 6.3 | 3309400 | 10 | US-09-738-626-1      | Sequence 1, Appli |
| 20 | 83.2  | 6.2 | 415     | 11 | US-09-918-995-17742  | Sequence 17742, A |
| 21 | 73.4  | 5.5 | 888     | 10 | US-09-726-397A-4     | Sequence 4, Appli |
| 22 | 73.4  | 5.5 | 888     | 10 | US-09-712-363-28     | Sequence 28, Appl |
| 23 | 73.4  | 5.5 | 86114   | 15 | US-10-080-170-648    | Sequence 648, App |
| 24 | 72.8  | 5.4 | 931     | 13 | US-10-027-632-161469 | Sequence 161469,  |
| 25 | 72.8  | 5.4 | 931     | 13 | US-10-027-632-161470 | Sequence 161470,  |
| 26 | 70.6  | 5.3 | 780     | 10 | US-09-974-300-343    | Sequence 343, App |
| 27 | 69.6  | 5.2 | 843     | 13 | US-10-027-632-170626 | Sequence 170626,  |
| 28 | 62    | 4.6 | 855     | 14 | US-10-128-714-2252   | Sequence 2252, Ap |
| 29 | 62    | 4.6 | 855     | 14 | US-10-128-714-7252   | Sequence 7252, Ap |
| 30 | 59.8  | 4.5 | 1014    | 10 | US-09-070-927A-157   | Sequence 157, App |
| 31 | 54.6  | 4.1 | 2908    | 14 | US-10-128-714-252    | Sequence 252, App |
| 32 | 54.6  | 4.1 | 2908    | 14 | US-10-128-714-1252   | Sequence 1252, Ap |
| 33 | 54.4  | 4.1 | 908     | 14 | US-10-128-714-1252   | Sequence 6252, Ap |
| 34 | 54.4  | 4.1 | 908     | 14 | US-09-726-397A-2     | Sequence 2, Appli |
| 35 | 53.6  | 4.0 | 2209    | 10 | US-09-796-692-6261   | Sequence 6261, Ap |
| 36 | 52.2  | 3.9 | 334     | 14 | US-10-040-862-6261   | Sequence 6261, Ap |
| 37 | 52.2  | 3.9 | 334     | 14 | US-10-156-761-390    | Sequence 390, App |
| 38 | 48.2  | 3.6 | 2817    | 14 | US-10-156-761-1531   | Sequence 1531, Ap |
| 39 | 47.4  | 3.5 | 888     | 14 | US-10-156-761-1      | Sequence 1, Appli |
| 40 | 47.4  | 3.5 | 9025608 | 14 | US-10-156-761-1729   | Sequence 1729, Ap |
| 41 | 47.2  | 3.5 | 795     | 14 | US-10-156-761-6088   | Sequence 6088, Ap |
| 42 | 46.4  | 3.5 | 1359    | 14 | US-10-156-761-1113   | Sequence 1113, Ap |
| 43 | 45.6  | 3.4 | 834     | 14 | US-10-156-761-2474   | Sequence 2474, Ap |
| 44 | 44.4  | 3.3 | 903     | 14 | US-10-037-270-807    | Sequence 807, App |
| 45 | 44    | 3.3 | 1645    | 14 |                      |                   |

ALIGNMENTS

RESULT 1  
US-10-161-418A-2  
; Sequence 2, Application US/10161418A  
; Publication No. US20030036078A1  
; GENERAL INFORMATION:  
; APPLICANT: EXELIXIS, INC.  
; TITLE OF INVENTION: P5CRS AS MODIFIERS OF THE P53 PATHWAY AND METHODS OF USE  
; FILE REFERENCE: EX02-055  
; CURRENT APPLICATION NUMBER: US/10/161.418A  
; CURRENT FILING DATE: 2002-09-11  
; PRIOR APPLICATION NUMBER: US 60/296,080  
; PRIOR FILING DATE: 2001-06-05  
; PRIOR APPLICATION NUMBER: US 60/328,509  
; PRIOR FILING DATE: 2001-10-10  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 2331  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-161-418A-2

|                       |              |  |                |              |
|-----------------------|--------------|--|----------------|--------------|
| Query Match           | 97.7%        | Score 1307;  | DB 14;         | Length 2331; |
| Best Local Similarity | 99.2%;       | Pred. No. 0;   |                |              |
| Matches 1313;         | Conservative | 0;   | Mismatches 10; | Indels 0;    |
| Gaps                  | 0;           |  |                |              |
| QY                    | 15           | GTCCGAGGCACACAGATGGCAGCTGCGGCGCGCTTCGCGCGCGTTCGTTGGGC    | 74             |              |
| Db                    | 12           | GTCCGAGGCACACAGATGGCAGCTGCGGCGCGCTTCGCGCGCGTTCGTTGGGC    | 71             |              |
| QY                    | 75           | CGGGCCCGCATGGCGGGGCCATCGCAGGCGCTCATCAGAGCAGGAAAAGTGAAGCT | 134            |              |
| Db                    | 72           | CGGGCCCGCATGGCGGGGCCATCGCAGGCGCTCATCAGAGCAGGAAAAGTGAAGCT | 131            |              |
| QY                    | 135          | CAGCACATCTGGCCAGTGCACACAGACAGGACCTATCTCACTTTCAAGCTCTGGGT | 194            |              |
| Db                    | 132          | CAGCACATCTGGCCAGTGCACACAGACAGGACCTATCTCACTTTCAAGCTCTGGGT | 191            |              |
| QY                    | 195          | TGCCGACACAGCATCCCAACAGGAGGTGTCGAGAGCTGCCTCTCTCTCTTTC     | 254            |              |





|                       |                 |   |            |              |
|-----------------------|-----------------|---|------------|--------------|
| Query Match           | 16.9%;          | Score 226.4;  | DB 14;     | Length 1848; |
| Best Local Similarity | 58.1%;          | Pred. No. 1.6e-51;  |            |              |
| Matches 467;          | Conservative 0; | Mismatches 316;   | Indels 21; | Gaps 3;      |
| QY                    | 56              | GGCGCTGGGCTTCGTGGCGGGCGCGCATCGCGGGGGCCATCGCGAGGGCCTCATCAG       | 115        |              |
| Db                    |                 |   |            |              |
| QY                    | 52              | GAGCGTGGGCTTCATCGGGCGCTGGCGAGCTGGCTTTGGCCCTGGCCAAAGGGCTTCACAGC  | 111        |              |
| Db                    |                 |   |            |              |
| QY                    | 116             | AGCAGAAAGTGGAAAGTCAAGCACATACTCGCCAGTGCACCAACACAGACAGAAACCTATG   | 175        |              |
| Db                    |                 |   |            |              |
| QY                    | 112             | AGCAGGCGCTCTGGCTGCCCAAGATAATGGCTAGCTCCC---CAGACATGGACCTGGC      | 168        |              |
| Db                    |                 |   |            |              |
| QY                    | 176             | TCACTTTCAAGTC-----TGGGTTGCCGAGCACGACGACATCCCAACAGAGAGTGCT       | 226        |              |
| Db                    |                 |   |            |              |
| QY                    | 169             | CACAGTTTCTGCTTCAGGAAGATGGGGGTGAAGTTGACACCCACACAAAGAGACGGT       | 222        |              |
| Db                    |                 |   |            |              |
| QY                    | 227             | GCAGAGCTGGCCTGCTCGTCACTTTTGCCACCAAGCCTCATGTGCTGCCAGCTGTCCTGGC   | 286        |              |
| Db                    |                 |   |            |              |
| QY                    | 229             | GCAGCACAGTATGTCTCTTCCTGGCTGTGAAGCCACACATCATCCCTTCATCTCTGGGA     | 288        |              |
| Db                    |                 |   |            |              |
| QY                    | 287             | AGAGTGGCTCCTGTGGTCCACACTGAACACATCTTGTGTGCCGTGCTGGGGTGTCTC       | 346        |              |
| Db                    |                 |   |            |              |
| QY                    | 289             | TGAATAGCGGCCACATTTGAGGACAGACACATTTGTGTGTCTGTGCGCGCGCGCTCAC      | 348        |              |
| Db                    |                 |   |            |              |
| QY                    | 347             | TCAGACACCTGGAGGAGCTGCTG-----CCGCCAACACACGGGTGCTGCGGGT           | 397        |              |
| Db                    |                 |   |            |              |
| QY                    | 349             | CATCAGCTCCATTCAGAGAAGCTGTGACGGCTTTTCGGCCACGCCCGCAGGCTCAATCCGCTG | 408        |              |
| Db                    |                 |   |            |              |
| QY                    | 398             | CTTGCCCAACCTGCCCTGTGTGTGTCCAGGAAGGGGCCATAGTATGCGCGGGGCGCCCA     | 457        |              |
| Db                    |                 |   |            |              |
| QY                    | 409             | CATGACCAACACTCCAGTCTGTGTGCGGGAGGGGCCACCGTGTATCCACAGGCAGCGCA     | 468        |              |
| Db                    |                 |   |            |              |
| QY                    | 458             | CGTGGGAGCAGCAGCACCAGCTTCCTGCAGCATCTCTGCTGAGGCCCTGTGGGCGGTGTGA   | 517        |              |
| Db                    |                 |   |            |              |
| QY                    | 469             | CGCCAGGTGGAGCAGCGGAGGCTCATGAGACAGCTGCTGAGCAGCGTGGGCTTCTGCAC     | 528        |              |
| Db                    |                 |   |            |              |
| QY                    | 518             | GGAGTGCCTGAAGCCTACGTCGACATCCACATGGCCTCAGTGGCAGTGGCGTGGCGCTT     | 577        |              |
| Db                    |                 |   |            |              |
| QY                    | 529             | GGAGGTGGAAGGAGCACCTGATTGATCGCTCACGGGCTCAGTGGCAGCGGCCCGGCTA      | 588        |              |
| Db                    |                 |   |            |              |
| QY                    | 578             | CGTGTGTGCATTCTCCAGGCGCCTGGCTGAAGAGCGCGTCAGATGGGCATGCCAGCAG      | 637        |              |
| Db                    |                 |   |            |              |
| QY                    | 589             | CGCATTCACAGCCTGTGATCCCTTGGCTGTATGGGGCGTGAAGATGGGACTTCCAAGCGC    | 648        |              |
| Db                    |                 |   |            |              |
| QY                    | 638             | CCTGGCCACCGCATCGCTGCCAGACCCCTGCTGGGACGCGCAAGATGCTGTGCACGA       | 697        |              |
| Db                    |                 |   |            |              |
| QY                    | 649             | CCTGGCAGTCCGCTCGGGGCCACAGGCCCTCTCTGGGGCTGCCAAGATGCTGTGCATC      | 708        |              |
| Db                    |                 |   |            |              |
| QY                    | 698             | GGGCCAACCCACCCAGCTGCGCTCAGACGTTGTGACCCGGGTGGCACCACTCTA          | 757        |              |
| Db                    |                 |   |            |              |
| QY                    | 709             | AGAACAGCACCCAGGCCAGCTCAAGGACAAACGTCAGCTCTCTCTGGTGGGCCACCATCCA   | 768        |              |
| Db                    |                 |   |            |              |
| QY                    | 758             | TGGACTCCAGCCCTTGAGCAGGCGGGCTGCCAGGACCCACCATGAGCGCGTGGAGGC       | 817        |              |
| Db                    |                 |   |            |              |
| QY                    | 769             | TGCTTGCAATGCTGGAGAGTGGGGCTTCCGCTCCCTGCTCATCAACGCTGTGGAGC        | 828        |              |
| Db                    |                 |   |            |              |
| QY                    | 818             | TGCCACTCGCGGGCCAAAGGAGCT  | 841        |              |
| Db                    |                 |   |            |              |
| QY                    | 829             | CTCTGTGATCCGCACACGCGGAGCT                                       | 852        |              |
| Db                    |                 |   |            |              |

## RESULT 5

```

US-09-880-107-2400
; Sequence 2400, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Wockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107

```

```

; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2400
; LENGTH: 1792
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 M77836
US-09-880-107-2400

```

| Query Match           | 16.8%;          | Score 224.8;   | DB 10;   | Length 1792; |
|-----------------------|-----------------|--|--|--------------|
| Best Local Similarity | 58.0%;          | Prod. No. 4.2e-51;   |  |              |
| Matches 466;          | Conservative 0; | Mismatches 317;  | Indels 21;   | Gaps 3;      |
| QY                    | 56              | GC   | CGCTGGGCTTCGTGGCGGGCGCCGATCGCGGGGGCCATCGCGAGGGCCCTCATCAG | 115          |
| DB                    | 14              | GAGCGTGGGCTTCATCGCGCTGCGCAGCTGGCTTTTGCCCTGCCAAGGGCTTCACAGC     | 73   |              |
| QY                    | 116             | AGCAGGAAAGTGGAAAGCTCAGCACATCTGGCCAGTGCACCAACAGACAGAGAACTATG    | 175  |              |
| DB                    | 74              | AGCAGGGCGTCTTGCTGCCCAAGATAATGGCTAGCTCC--CAGACATGGACCTGGC       | 130  |              |
| QY                    | 176             | TCACCTTCAAGCTC-----TGGGTGCGCGACACGACATCCACACAGGAGGTGCT         | 226  |              |
| DB                    | 131             | CACAGTTTCCTCTCAGGAAGATGGGGTGAAGTTGACACCCACCAAGAGAGACGGT        | 190  |              |
| QY                    | 227             | GCAGAGCTGGCTGCTCGTCTATCTTTGGCCACCAAGCCCTCATGTGTGGCAGCTGTCCTGGC | 286  |              |
| DB                    | 191             | GCACACAGTGAATGTGCTCTCTCTGGCTGTGAAGCCACACATATCCCTTCATCCTGGA     | 250  |              |
| QY                    | 287             | AGAGTGGCTCCTGTGTGCTCAACACTGAACACATCTTGGTGTCCGTGCTGCTGGGGTGTG   | 346  |              |
| DB                    | 251             | TGAATAGGCGCCGACATTGAGGACACACATTTGGTGTGCTGCGCGCGCGCGCTAC        | 310  |              |
| QY                    | 347             | TCGTGAGCACCTGGAGGAGCTGCTG-----CCCCAACACACAGGCTGCTGCGGGT        | 397  |              |
| DB                    | 311             | CATCAGCTCCATTTGAGAAGAAGTGTGAGGGTTTCGGCCAGCGCCAGGGTTCATCCGCTG   | 370  |              |
| QY                    | 398             | CTTGCCCAACCTGCCCCCTGTGTGTCTGAGGAAGGGCCATAGTGTATGCGGGGGCCGCA    | 457  |              |
| DB                    | 371             | CATGACCAACACTCCAGTCTGTGTGCGGAGGGGGCCACCGTGTATGCCACAGACAGCA     | 430  |              |
| QY                    | 458             | CGTGGGAGCAGGAGACACCAAGCTCTGTGACATCTGTGAGAGCCCTGTGGGGGTGTGA     | 517  |              |
| DB                    | 431             | CGCCACAGGTGGAGGAGCGGGAGGCTCATGAGACAGCTGCTGAGCAGCGTGGGCTTCGCAC  | 490  |              |
| QY                    | 518             | GGAGGTGCCTGAAGCCTACGCTCGACATCCACACTGGCCTCAGTGGCAGTGGCGTGGCCTT  | 577  |              |
| DB                    | 491             | GGAGGTGGAAGAGGACCTGATTGATGCCGTACAGGGGCTCAGTGGCAGCGGCCCGCCCTA   | 550  |              |
| QY                    | 578             | CGTGTGTGATTTCTCGAGGCCCTGGCTGAAGAGCCGCTCAAGATGGGCTATGCCAGCAG    | 637  |              |
| DB                    | 551             | CGCATTCACAGCCCTGGATGCCCCCTGGCTGATGGGGTGTGAAGATGGGACTTCCAAGCG   | 610  |              |
| QY                    | 638             | CTTGGCCACCGCATCGCTGCCACAGACCTGCTGGGAGCGGCCCAAGATGCTGTGACACA    | 697  |              |
| DB                    | 611             | CCTGGCAGTCCGCTCGGGGCCAGGCCCTTCTTGGGGGCTGCCAAGATGCTGTGCTGACT    | 670  |              |
| QY                    | 698             | GGGCCAACCCAGCCAGCTGCGCTCAGACGCTGTGCACCCCGGGTGGCACCACTATCTA     | 757  |              |
| DB                    | 671             | AGAACAGCACCCAGGCGACTCAAGGACAAACGTCAGCTCTCTGGTGGGGCCACCATCCA    | 730  |              |
| QY                    | 758             | TGGACTTCCACGCCCTTGGAGCAGGGCGGCTGCGAGACGCCACCATCAGGCCCTGAGGC    | 817  |              |
| DB                    | 731             | TGCCCTTGATGTGCTGGAGAGTGGGGGCTTCCGCTCCCTGCTCATCAACGCTGTGAGGC    | 790  |              |
| QY                    | 818             | TGCCACCTCTCCGGGGCAAGGAGCT                                      | 841  |              |

```
Db      791 CTCCTGCATCCGCACACGGGAGCT 814
RESULT 6
US-10-161-418A-1
; Sequence 1, Application US/10161418A
; Publication No. US20030036078A1
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: P5CRS AS MODIFIERS OF THE P53 PATHWAY AND METHODS OF USE
; FILE REFERENCE: EX02-055
; CURRENT FILING DATE: 2002-09-11
; PRIOR FILING DATE: 2002-09-11
; PRIOR APPLICATION NUMBER: US 60/296,080
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/328,509
; PRIOR FILING DATE: 2001-10-10
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1792
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-161-418A-1
Query Match      16.8%; Score 224.8; DB 14; Length 1792;
Best Local Similarity 58.0%; Pred. No. 4.2e-51;
Matches 466; Conservative 0; Mismatches 317; Indels 21; Gaps 3;
Qy      56 GCGCGTGGGCTTGGTGGGCGCGGCGCATGGCGGGGCCATCGCGAGGGCCTCATCAG 115
Db      14 GAGCGTGGGCTTATCGCGCGCTGGCCAGCTGTTTGGCTTGGCCAGGGGCTTCACAGC 73
Qy      116 AGCAGGAAAGTGGAGCTCAGCAGACATCTGCGCAGTGCACCAACAGACAGAACCTATG 175
Db      74 AGCAGGCGTGTGGTGGTGGCCCAAGATATGGCTAGCTCCC---CAGACATGGAGCTGGC 130
Qy      176 TCACCTTTCAAGCTC-----TGGGTTGCCGACCGCAGCTCTCCAAACAGGAGTGCT 226
Db      131 CACAGTTCTGCTCTCAGGAAGATGGGGTGAAGTTGACACCCCAACAGAGACGCT 190
Qy      227 GCAGAGTGCGCTCGTCTCATCTTTGCCACCAAGCCTCATGTGCTGCCAGCTGTCTGCGC 286
Db      191 GCAGCAGTGTGTCTTCTTCTGCTGTGAAGCCACACATCATCCCCCTTCATCCGTGA 250
Qy      287 AGAGTGGCTCCGTGTGCTCACCCTGAACACATCTTGGTGTGCTGCTGCTGGGCTGTC 346
Db      251 TGAATAGGCGCGACATTTGAGACAGACATTTGGTGTCTTGGCGCGCGCGGTGCAC 310
Qy      347 TCTGAGCAGCCTGGAGGAGCTGCTG-----CCCCAAACACACGGGTGCTGCGGCT 397
Db      311 CATCAGCTCCATTGAGAAGAGCTGTACAGCGTTTTCGGCCAGCGCCCGAGGCTATCGGCTG 370
Qy      398 CTTGCCCAACCTGCCCTGTGTGTCAGGAGGGGCCATGTGATGGCGGGGGCGGCCA 457
Db      371 CATGACCAACATCCAGTCTGTGCGGGAGGGGCCACCGGTATGCCACAGACGCA 430
Qy      458 CGTGGGAGCAGCAGACCAAGCTCCTGCAGCATCTGCTGGAGGCGCTGTGGGCGGTGA 517
Db      431 CGGCCAGGTGGAGAGCGGGAGGCTCATGGAGAGCTGTGAGCAGCGTGGGCTTCTGCAC 490
Qy      518 GGAGTGCCTTGAAGCGCTTACCTGCACATCCACATGGCCTGAGTGGCGAGTGGCGCTT 577
Db      491 GGAGTGGAGAGGACCTGATTGATCCGCTCAGCGGGCTCAGTGGCAGCGGCCCGCTCA 550
Qy      578 CGTGTGTGCAATTTCCGAGGCGCTGTGTAAGAGCGCGTCAAGATGGGCGATGCCAGCAG 637
Db      551 CGCATTCACAGCCCTGGATGCCCTGTGCTGATGGGGTGTGAAGATGGGACTTCCAAGGCG 610
Qy      638 CTTGGGCCACCGCATCGCTGCCAGACCCCTGCTGGGAGCGGCCCAAGATGCTGTGCACGA 697
Db      611 CTTGGGATCTCGGCTCGGGGCCCGAGGCCCTCTCTGGGGGTGCGCAAGATGCTGCTGCACTC 670
Qy      698 GGGCCAAACCCAGCCAGCCAGCTGCGCTCAGACGTGTGCACCCCGGGTGGCAGCCACTCTA 757
Db      671 AGAACAGACCCAGCCAGCCAGCTCAAGGACAACGTCAGCTCTCTGTTGGGGCCACCATCCA 730
Qy      758 TGGACTCACGCCCTTGGAGAGGGCGGGCTGCGAGCAGCCACCATGATGAGCCCGTGGAGGC 817
Db      731 TGCCTTGCATGTGCTGGAGAGTGGGGCTTCCGCTCCCTGCTCTCATCAACGCTGTGGAGGC 790
Qy      818 TGGCACCTTGGCGGCCCAAGGAGCT 841
Db      791 CTCCTGCATCCGCACACGGGAGCT 814
RESULT 7
US-10-161-418A-5
; Sequence 5, Application US/10161418A
; Publication No. US20030036078A1
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: P5CRS AS MODIFIERS OF THE P53 PATHWAY AND METHODS OF USE
; FILE REFERENCE: EX02-055
; CURRENT FILING DATE: 2002-09-11
; PRIOR FILING DATE: 2002-09-11
; PRIOR APPLICATION NUMBER: US 60/296,080
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/328,509
; PRIOR FILING DATE: 2001-10-10
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 1715
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-161-418A-5
Query Match      16.0%; Score 214.4; DB 14; Length 1715;
Best Local Similarity 58.0%; Pred. No. 2.7e-48;
Matches 466; Conservative 0; Mismatches 316; Indels 22; Gaps 4;
Qy      56 GCGCGTGGCTTTCGTGGGCGCGGCGCATGGCGGGGCCATCGCGCAGGCGCTCATCAG 115
Db      281 GAGCGTGGGCTTCATCGCGCGCTGGCCAGCTGGCTTTTCCCTTGGCCAGGCTTTTCACAGC 340
Qy      116 AGCAGGAAAGTGGAGCTCAGCAGACATCTGCGCAGTGCACCAACAGACAGAACCTATG 175
Db      341 AGCAGGCGTCTTGGCTGCGCCACAGATAATGGCTAGCTCCC---CAGACATGGAGCTGCGC 397
Qy      176 TCACCTTTCAAGCTC-----TGGGTTGCCGACCGCAGCTCTCCAAACAGGAGGTGCT 226
Db      398 CACAGTTTCTGCTCTCAGGAAGATGGGGTGAAGTTGACACCCCAACAGAGGACCGGT 457
Qy      227 GCAGAGCTGCCCTGCTCGTCATCTTTGCCACCAAGCCTCATGTGCTGCCAGCTGTCTGCGC 286
Db      458 GCAGCAGACATGATGTGCTCTTCTGCTGTGAGGCCACACATCATCCCTTCATCCTGGA 517
Qy      287 AGAGTGGCTCCTGTGGTGCACCATGTGAACACATCTTGGTGTGCTGCTGCTGGGCTGTC 346
Db      518 TCAATAGCGCGCGACATTTGAGGACAGACACATTTGGTGTCTGCTGCGGGCGCGGCTCAC 577
Qy      347 TCTGAGCAGCCTGGAGGAGCTGCTG-----CCCCAAACACACGGGTGCTGCGGCT 397
Db      578 CATCAGCTCCATTGAGAAGAGCTGTACAGCGTTTTCGGCCAGCGCCCGAGGCTCATCCGCTG 637
Qy      398 CTTGCCCAACCTGCCCTGTGTGTCAGGAAGGGCCATAGTGTGGGCGGGCGGCCA 457
Db      638 CATGACCAACATCCAGTCTGTGCGGGAGGGGCCACCGCTGTATGCCACAGGACGAC 697
Qy      458 CGTGGGAGCAGCAGACCAAGCTCCTGCAGCATCTGCTGGAGGCGCTGTGGGCGGTGTA 517
Db      698 GCCCAGGTGGAG-GACGGGAGGCTCATGGAGAGCTGTGAGCAGCGTGGGCTTCTGTCAC 756
Qy      518 GGAGTGCCTGAAGCGCTAGCTGCAGATCCACACTGGCCTCAGTGGCAGTGGGCTGGGCTT 577
Db      518 GGAGTGCCTGAAGCGCTAGCTGCAGATCCACACTGGCCTCAGTGGCAGTGGGCTGGGCTT 577
```

Db 757 GGAGGTGAAGAGAGACCTGATTGATGCCGTACAGGGGCTCAGTGGCAGGGGCCCGCCCTTA 816  
Qy 578 CGTGTGTGCTATCCAGAGCCCTGGCTGAAGAGCGGTCAAGATGGCATGCCAGCAG 637  
Db 817 CGCATTACAGCCCTGATGCCCTGGCTGATGGGGTGTGAAGATGGGACTTCCAAGGCG 876  
Qy 638 CTTGGCCCAACCGCATCGCTCCCAAGACCCCTGCTGGGAGGGCCAAAGATGCTGTGCACGA 697  
Db 877 CTTGGCAGTCCGCTCGGGGCCCAAGGCCCTCTCTGGGGGTGCGCAAGATGCTGTGCACTC 936  
Qy 698 GGCCCAACACCCAGCCAGCTGGCTCAGACGTGTGCACCGGTGTGCACCCCGGTGGCACCACTCTA 757  
Db 937 AGAACAGACCCAGGCGGTCAAGGACAACTGCTCTCTGGTGGGGCCACCATCA 996  
Qy 758 TGGACTCCAGCCCTGGAGCAGGGCGGTGCGAGCAGCACCATGAGCGCGGTGGAGGC 817  
Db 997 TGCCTTCATGTCTGGAGAGTGGGGCTTCCTGCTCCCTGCTCATCAACGCTGTGGAGGC 1056  
Qy 818 TGGCACTGCCGGGCCAAGAGCT 841  
Db 1057 CTCTGCTATCCGACACAGGAGCT 1080

## RESULT 8

US-10-161-418A-4  
; Sequence 4, Application US/10161418A  
; Publication No. US20030036078A1  
; GENERAL INFORMATION:  
; APPLICANT: EXELIXIS, INC.  
; TITLE OF INVENTION: P5CRs AS MODIFIERS OF THE P53 PATHWAY AND METHODS OF USE  
; FILE REFERENCE: EX02-055  
; CURRENT APPLICATION NUMBER: US/10/161,418A  
; CURRENT FILING DATE: 2002-09-11  
; PRIOR APPLICATION NUMBER: US 60/296,080  
; PRIOR FILING DATE: 2001-06-05  
; PRIOR APPLICATION NUMBER: US 60/328,509  
; PRIOR FILING DATE: 2001-10-10  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 4  
; LENGTH: 1028  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (999)..(999)  
; OTHER INFORMATION: "n" is A, C, G, or T  
US-10-161-418A-4

Query Match 15.8%; Score 211.2; DB 14; Length 1028;  
Best Local Similarity 58.2%; Pred. No. 1.9e-47;  
Matches 438; Conservative 0; Mismatches 293; Indels 21; Gaps 3;  
Qy 56 GCGCGTGGCTTCGTGGGCGCGCGCATGCGGGGCGCATCGCGAGGCGCTCATCAG 115  
Db 61 GAGCGTGGCTTCATCGGGCGTGGCGAGCTGCTTTGCGCTGGCAGGCGTTCACAGC 120  
Qy 116 AGCAGGAAGTGGAGCTCAGCATATCTAGTGGCAGTGACCAACAGACAGAACTATG 175  
Db 121 AGCAGGCGTCTTGGTGGTGGCCCAAGATAATGCTAGCTCC--CAGACATGGACCTGGC 177  
Qy 176 TCACATTTCAGCTC-----TGGGTTGGCGGACACGACCTCCCAACAGGAGGTGCT 226  
Db 178 CACAGTTTCTCTCAGGAAGATGGGGTGAAGTTGACCCCAACAGAGGAGCGGT 237  
Qy 227 GCAGAGCTGCCGTGCTGTCATCTTTGCCACCAAGCCTCATGTGTCGACAGTGTCTGGC 286  
Db 238 GCAGCAGAGTATGCTGCTCTTCTGCTGTGAAGCCACATCATCCCTTCATCCTTGA 297  
Qy 287 ACAGTGGCTCTGCTGGTCCACCACTGACACATCTTGTGTCGCTGGGTGCTGGGGTGC 346  
Db 298 TGAATAGCCCGCCACATTGAGGACAGACATATGTTGTTGCTCTGCGCGCGCGGCTCAC 357

Qy 347 TCTGAGCACCTGGAGGAGCTGCTG-----CCCCCAACACACGCGTGTCTGCGGGT 397  
Db 358 CATCAGCTCCATTGAGAAGAAGCTGTACGCGTTTGGCCAGCCCGCCAGGGTTCATCCGCTG 417  
Qy 398 CTTGCCCAACCTGCCCTGTGTGGTCCAGGAAGGGCCATAGTGTATGGCGGGGCGCGCA 457  
Db 418 CATGACCAACACTCCAGTCTGTGGCGGAGGGGCGCACCGTGTATGCCACAGCAGCA 477  
Qy 458 CTTGGGAGCAGCAGAGACCAAGCTCTCTCAGCATCTCTGTGAGGCCCTGTGGGCGGTGTGA 517  
Db 478 GCGCCAGTGGAGGACGGGAGGCTCATGGAGCAGCTGTGAGCAGCGTGGGCTTCTGCAC 537  
Qy 518 GGAGTGTCTCAAGCCTACGTCGACATCCACACTGGCTCAGTGGCTCAGTGGCGTGGCTG 577  
Db 538 GGAGTGAAGAGACCTGATGATGCGCTCACGGGCTCAGTGGCAGCGGCCCGCCTA 597  
Qy 578 CGTGTGTGCTTCTCCAGGCCCTGGCTGAAGAGCGGTCAAGATGGCGCATGCCAGCAG 637  
Db 598 CGCATTCACAGCCCTGGATGCCCTGGCTGATGGGGTGTGAAGATGGGACTTCCAAAGCG 657  
Qy 638 CTTGGCCCAACCGCATCGCTGCCAGACCCCTGCTGGGAGCGGCCAAGATGCTGTGCACGA 697  
Db 658 CTTGGCAGTCCGCTCGGGGCGCGAGGCCCTCTCTGGGGCTGCCAAGATGCTGTGCAC 717  
Qy 698 GGGCAACACCCAGCGCGCTGCGCTCAGACGTGTGCACCCCGGGTGGCAGCACCATCTA 757  
Db 718 AGAACAGCACCAGCGCAGCTCAAGGACAGGTACGTCTCTGCTGGGCGCACCATCCA 777  
Qy 758 TGGACTCCAGCCCTGGAGCAGGCGGGCTGC 789  
Db 778 TGCCTTCATGTCTGGAGAAGTGGGGCTTC 809

## RESULT 9

US-09-925-300-278  
; Sequence 278, Application US/09925300  
; Patent No. US20020151681A1  
; GENERAL INFORMATION:  
; APPLICANT: Craig Rosen,  
; APPLICANT: Steve Ruben  
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
; FILE REFERENCE: PA101  
; CURRENT APPLICATION NUMBER: US/09/925,300  
; CURRENT FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: PCT/US00/05988  
; PRIOR FILING DATE: 2000-03-08  
; PRIOR APPLICATION NUMBER: 60/124,270  
; PRIOR FILING DATE: 1999-03-12  
; NUMBER OF SEQ ID NOS: 1890  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 278  
; LENGTH: 1478  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-925-300-278

Query Match 15.2%; Score 203.4; DB 10; Length 1478;  
Best Local Similarity 57.9%; Pred. No. 2.5e-45;  
Matches 466; Conservative 0; Mismatches 316; Indels 23; Gaps 5;  
Qy 56 GCGCGTGGGCTTCGTGGGCGCGCGCATGCGGGGCGCATCGCGAGGCGCTCATCAG 115  
Db 78 GAGCGTGGGCTTCATCGGCGCTGGCCAGCTGGCTTTTGGCCTGGCCAAAGGCTTGCACAG 137  
Qy 116 -AGCAGGAAAGTGGAGCTCAGCATACTGCGCCAGTCACCAACACAGACAGAACTAT 174  
Db 138 CAGCAGCGCTCTGGCTGCCCAAGATATGCTAGCTCCC---CAGACATGGACCTGG 194  
Qy 175 GTCACTTTCAAGCTC-----TGGGTTGGCGGACCGACGCTCCCAACAGGAGGTGC 225  
Db 195 CCACAGTTTCTGCTCTCAGGAAGATGGGGTGAAGTTGACACCCCAACAAAGGAGCGG 254  
Qy 226 TGCAGAGCTGCCCTGCTCGTCACTTTTGGCAACCAAGCTCATGTGCTGCCAGCTGTCTCTCG 285



| Query Match           | 15.1k | Score   | 201.8 | DB         | 9   | Length | 1742 |
|-----------------------|-------|---|-------|------------|-----|--------|------|
| Best Local Similarity | 55.0k | Pred.   | No.   | 6.9e-45    |     |        |      |
| Matches               | 448   | Conservative  | 0     | Mismatches | 352 | Indels | 15   |
|                       |       |   |       |            |     |        | Gaps |
| QY                    | 42    | GAGCCGCTCTCGCGGCGCTGGCTGCTTCGTGGCGGGGGCCGCGCATGSCGGGGGCCATCGC | 101   |            |     |        |      |
| Db                    | 180   | GAACCGCGACCATGACGCTGGCGCTTCATCGGGCGCGCCAGCTGGCTATGCTCTGGCG    | 239   |            |     |        |      |
| QY                    | 102   | CAGGGCTCTCATGAGACGAAAGTGGAGCTGTGACGACATACTGGCCAGTGCACCAACA    | 161   |            |     |        |      |
| Db                    | 240   | CGGGGCTTCACGGCCGACGACATCCCTGTGCGGTCTCAAGATATAGCCAGCTCCCCGAA   | 299   |            |     |        |      |
| QY                    | 162   | GACAGGAACC-----TATGTCACTTTCAAGCTCTGGGTTGCCGACACGACACTCCAAC    | 215   |            |     |        |      |
| Db                    | 300   | ATGAACTGCCACCGGTGCCGGCTCAGGAAGATGGGTGTGAACCTGCACAGCAGCAAC     | 359   |            |     |        |      |
| QY                    | 216   | CAGGAGTGTGCAGAGCTGCCCTGCTGCTCATCTTTGCCAACCAAGCCCTCATGTGCTGCCA | 275   |            |     |        |      |
| Db                    | 360   | AAGGAGACGGTGAAGCACACGACAGCTCCTGTTCTGGCTGTGAAGCCACATATCATCCCC  | 419   |            |     |        |      |
| QY                    | 276   | GCTGCTCTGSCAGAGTGGCTCCTGTGGTCACCACCTGACACATCTTGTTGCTCGGTGGCT  | 335   |            |     |        |      |
| Db                    | 420   | TTCACTCTGATGAGATTGGGCGCAGCTGCAGCCAGACACATCGTGGTCTCTCTGTGGCG   | 479   |            |     |        |      |
| QY                    | 336   | GCTGGGTGTCTCTGACACCCCTGGAGAGCTGCTG-----CCCCAAACACACAGG        | 386   |            |     |        |      |
| Db                    | 480   | GCTGGTGTCAACATCAGCTCTGTGAGAGAAGAGCTGATGGCATTCACGCCAGCCCCAAA   | 539   |            |     |        |      |
| QY                    | 387   | GTGCTGCGGGTCTTGCCCAACCTCGCTGTGGTTCAGGAGGGCCATAGTGNATGGCG      | 446   |            |     |        |      |
| Db                    | 540   | GTGATTCGCTCATGACCAACACACCTGTGGTAGTCAGGAAGGCGCTACAGCTGTACGCC   | 599   |            |     |        |      |
| QY                    | 447   | CGGGGCCGCCACCTGGGGAGCAGCAGACCAAGCTCCTGCACATCTGCTGGAGGCGCTGT   | 506   |            |     |        |      |
| Db                    | 600   | ACGGGCACCATGCCCTGGTGGAGGATGGCGAGCTCCTGGAGCAGCTCATGACGAGCGTG   | 659   |            |     |        |      |
| QY                    | 507   | GGGCGGTGTGAGGAGTGCCTGAAGCCTACGCTGCACATCCACATGCGCCTCAGTGGCAGT  | 566   |            |     |        |      |
| Db                    | 660   | GGCTTCGCACTGAGGTGGAAGAGGACCTCATCTGATGGCGCTCAAGGGGCTCAGTGGCAGC | 719   |            |     |        |      |
| QY                    | 567   | GGCTGGCCCTCGTGTGTCATCTCCGAGGCGCCTGGCTGAAGGAGCGCTCAAGATGGC     | 626   |            |     |        |      |







GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 8, 2003, 14:36:31 ; Search time 396 Seconds  
(without alignments)  
9120.829 Million cell updates/sec

Title: US-09-806-536A-29  
Perfect score: 1338  
Sequence: 1 ggtgagcgcagtcttcga.....ggggcagcagcatatggggt 1338

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

| Database : | N_Geneseq_19Jun03.*                                |
|------------|--|
| 1:         | /SIDSL1/gcgdata/geneseq/geneseq-nbml/NA1980.DAT.*  |
| 2:         | /SIDSL1/gcgdata/geneseq/geneseq-nbml/NA1981.DAT.*  |
| 3:         | /SIDSL1/gcgdata/geneseq/geneseq-nbml/NA1982.DAT.*  |
| 4:         | /SIDSL1/gcgdata/geneseq/geneseq-nbml/NA1983.DAT.*  |
| 5:         | /SIDSL1/gcgdata/geneseq/geneseq-nbml/NA1984.DAT.*  |
| 6:         | /SIDSL1/gcgdata/geneseq/geneseq-nbml/NA1985.DAT.*  |
| 7:         | /SIDSL1/gcgdata/geneseq/geneseq-nbml/NA1986.DAT.*  |
| 8:         | /SIDSL1/gcgdata/geneseq/geneseq-nbml/NA1987.DAT.*  |
| 9:         | /SIDSL1/gcgdata/geneseq/geneseq-nbml/NA1988.DAT.*  |
| 10:        | /SIDSL1/gcgdata/geneseq/geneseq-nbml/NA1989.DAT.*  |
| 11:        | /SIDSL1/gcgdata/geneseq/geneseq-nbml/NA1990.DAT.*  |
| 12:        | /SIDSL1/gcgdata/geneseq/geneseq-nbml/NA1991.DAT.*  |
| 13:        | /SIDSL1/gcgdata/geneseq/geneseq-nbml/NA1992.DAT.*  |
| 14:        | /SIDSL1/gcgdata/geneseq/geneseq-nbml/NA1993.DAT.*  |
| 15:        | /SIDSL1/gcgdata/geneseq/geneseq-nbml/NA1994.DAT.*  |
| 16:        | /SIDSL1/gcgdata/geneseq/geneseq-nbml/NA1995.DAT.*  |
| 17:        | /SIDSL1/gcgdata/geneseq/geneseq-nbml/NA1996.DAT.*  |
| 18:        | /SIDSL1/gcgdata/geneseq/geneseq-nbml/NA1997.DAT.*  |
| 19:        | /SIDSL1/gcgdata/geneseq/geneseq-nbml/NA1998.DAT.*  |
| 20:        | /SIDSL1/gcgdata/geneseq/geneseq-nbml/NA1999.DAT.*  |
| 21:        | /SIDSL1/gcgdata/geneseq/geneseq-nbml/NA2000.DAT.*  |
| 22:        | /SIDSL1/gcgdata/geneseq/geneseq-nbml/NA2001A.DAT.* |
| 23:        | /SIDSL1/gcgdata/geneseq/geneseq-nbml/NA2001B.DAT.* |
| 24:        | /SIDSL1/gcgdata/geneseq/geneseq-nbml/NA2002.DAT.*  |
| 25:        | /SIDSL1/gcgdata/geneseq/geneseq-nbml/NA2003.DAT.*  |

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score  | Query Match | Length | DB ID | Description                 |
|------------|--------|-------------|--------|-------|-----------------------------|
| 1          | 1338   | 100.0       | 1338   | 21    | AAA09388 Human DNA encoding |
| 2          | 1307   | 97.7        | 2331   | 22    | AAH18288 Human cDNA sequenc |
| 3          | 1307   | 97.7        | 2331   | 25    | ABA00759 P5CR related nucle |
| 4          | 1302.2 | 97.3        | 2496   | 22    | AAH77597 Human dihydropyrr  |
| 5          | 1291.8 | 96.5        | 2615   | 22    | AAH14370 Human cDNA sequenc |
| 6          | 1127.8 | 84.3        | 1178   | 25    | ABA00764 P5CR related nucle |
| 7          | 857    | 64.1        | 860    | 25    | ABA00763 P5CR related nucle |
| 8          | 737.6  | 55.1        | 794    | 22    | AAH08500 Human cDNA clone ( |

|    |       |      |         |    |                             |
|----|-------|------|---------|----|-----------------------------|
| 9  | 723   | 54.0 | 5055    | 22 | AAK51999 Human polynucleoti |
| 10 | 599.4 | 44.8 | 5647    | 22 | AA542169 Genomic sequence # |
| 11 | 486.6 | 36.4 | 563     | 22 | AAH06264 Human cDNA clone ( |
| 12 | 241.8 | 18.1 | 270     | 22 | AA541342 cDNA encoding nove |
| 13 | 226.4 | 16.9 | 1848    | 25 | ABA00760 P5CR related nucle |
| 14 | 224.8 | 16.8 | 1792    | 24 | ABT10919 Human breast cance |
| 15 | 224.8 | 16.8 | 1792    | 24 | ABN95903 Gene #2401 used to |
| 16 | 224.8 | 16.8 | 1792    | 25 | ABA00758 P5CR related nucle |
| 17 | 224.8 | 16.8 | 1889    | 24 | ABO54355 Human ovarian anti |
| 18 | 214.4 | 16.0 | 1715    | 25 | ABA00762 P5CR related nucle |
| 19 | 211.2 | 15.8 | 1028    | 25 | ABA00761 P5CR related nucle |
| 20 | 203.4 | 15.2 | 1478    | 21 | AAF15843 Human prostate can |
| 21 | 201.8 | 15.1 | 1708    | 25 | ABA00765 P5CR related nucle |
| 22 | 201.8 | 15.1 | 1742    | 21 | AAA88073 Human delta 1-pyrr |
| 23 | 201.8 | 15.1 | 1742    | 22 | AAD20894 Human delta 1-pyrr |
| 24 | 201.8 | 15.1 | 1742    | 24 | ABK87030 Human Py-CR encodi |
| 25 | 197   | 14.7 | 1121    | 22 | AAF81847 cDNA encoding nove |
| 26 | 194.4 | 14.5 | 251     | 22 | AA541363 Human dithp enzyme |
| 27 | 183.6 | 13.7 | 1587    | 25 | ACC46089 P5CR related nucle |
| 28 | 167   | 12.5 | 999     | 25 | ABA00766 Human cDNA sequenc |
| 29 | 148.8 | 11.1 | 3149    | 22 | AAH14378 Drosophila melanog |
| 30 | 143   | 10.7 | 867     | 23 | ABL10077 Drosophila melanog |
| 31 | 127.6 | 9.5  | 3120    | 23 | ABL10076 Drosophila melanog |
| 32 | 127.6 | 9.5  | 4978    | 23 | ABL10046 Drosophila melanog |
| 33 | 118.2 | 8.8  | 1183    | 23 | ABL29275 Drosophila melanog |
| 34 | 112.4 | 8.4  | 1868    | 23 | AA584881 DNA encoding novel |
| 35 | 110.2 | 8.2  | 3244    | 23 | ABL29274 Drosophila melanog |
| 36 | 110.2 | 8.2  | 3284    | 23 | ABL29246 Drosophila melanog |
| 37 | 105   | 7.8  | 1163020 | 24 | ABQ67197 Listeria innocua c |
| 38 | 105   | 7.8  | 3011208 | 24 | ABQ69245 Listeria innocua c |
| 39 | 94.2  | 7.0  | 1837    | 22 | AAH48471 Escherichia coli p |
| 40 | 87.4  | 6.5  | 56050   | 23 | AA559549 Propionibacterium  |
| 41 | 84.4  | 6.3  | 810     | 22 | AAH65425 C glutamicum codin |
| 42 | 84.4  | 6.3  | 933     | 22 | AAF71901 Corynebacterium gl |
| 43 | 84.4  | 6.3  | 349980  | 22 | AAH68525 C glutamicum codin |
| 44 | 84.2  | 6.3  | 2944528 | 24 | ABA03041 Listeria monocytog |
| 45 | 82.6  | 6.2  | 6743    | 24 | ABQ71038 Listeria monocytog |

# ALIGNMENTS

|          |  |
|----------|--|
| RESULT 1 |  |
| AAA09388 |  |
| ID       | AAA09388 standard; DNA; 1338 BP.                                       |
| XX       | AAA09388;  |
| AC       |  |
| XX       | 10-AUG-2000 (first entry)  |
| DT       |  |
| XX       | Human DNA encoding OXRE-14.  |
| DE       |  |
| XX       | OXRE-14; oxidoreductase; pyrroline-5-carboxylate-reductase;            |
| KW       | antiproliferative; anticancer; hepatotropic; antiviral; antiasthmatic; |
| KW       | neurotropic; neuroprotectant; antiparkinsonian's; antisclerotic;       |
| KW       | anxiolytic; antischizophrenic; anti-inflammatory; antiarthritic; ss.   |
| OS       | Homo sapiens.  |
| XX       |  |
| FH       | Key Location/Qualifiers  |
| FT       | CDS 30..854  |
| FT       | /*tag= a   |
| FT       | /product= OXRE-14  |
| XX       |  |
| PN       | WO200020604-A2.  |
| PD       | 13-APR-2000.   |
| XX       |  |
| PF       | 06-OCT-1999; 99WO-US23434.   |
| XX       |  |
| PR       | 06-OCT-1998; 98US-0172227.   |
| PR       | 02-DEC-1998; 98US-0155202.   |
| PR       | 10-MAR-1999; 99US-0123911.   |

XX (INCY-) INCYTE PHARM INC.  
PA Lal P, Guegler KJ, Gorgone GA, Corley NC, Baughn MR, Tang YN;  
PI Hillman JL, Bandman O, Azimzai Y, Au-Young J, Yue H, Lu DAM;  
PI Yang J;  
XX WPI; 2000-303785/26.  
DR P-PSDB; AAY92517.  
XX Purified polypeptide for treating or preventing disorders associated  
PT with decreased expression or activity of oxidoreductase molecules  
XX Claim 9; Page 96; 97pp; English.  
XX AAA09375-89 encode oxidoreductases, designated OXRE-1 to -15. OXRE-14  
CC has identity with pyrroline-5-carboxylate-reductase.  
CC The polypeptides are useful for treating or preventing a disorder  
CC associated with decreased expression or activity of OXRE. Antagonists  
CC of OXRE are useful for treating or preventing a disorder associated with  
CC increased expression or activity of OXRE. The disorders include cell  
CC proliferative disorders (cirrhosis, hepatitis), cancer (leukemia,  
CC melanoma), hypopituitarism and hyperpituitarism, hypothyroidism and  
CC hyperthyroidism, metabolic disorders (Addison's disease, cystic  
CC fibrosis), reproductive disorders (infertility, ovulatory defects),  
CC neurological disorders (Alzheimer's disease, Parkinson's disease,  
CC multiple sclerosis), mental disorders (anxiety, schizophrenia),  
CC autoimmune/inflammatory disorders (acquired immunodeficiency syndrome  
CC (AIDS), asthma, osteoarthritis), and viral infections. The  
CC polynucleotides may be used in Southern or Northern analysis, polymerase  
CC chain reaction (PCR), or in enzyme-linked immunosorbent assays (ELISA).  
XX Sequence 1338 BP; 225 A; 423 C; 440 G; 250 T; 0 other;  
S  
Query Match 100.0%; Score 1338; DB 21; Length 1338;  
Best Local Similarity 100.0%; Pred. No. 3 5e-289;  
Matches 1338; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 GGTGAGCGCAGTCTGTCCGAGGCAACAAGATGCGAGTGGCGAGCGGTCTCCGCGCGCG 60  
Db 1 GGTGAGCGCAGTCTGTCCGAGGCAACAAGATGCGAGTGGCGAGCGGTCTCCGCGCGCG 60  
Qy 61 TGGGCTTCGTTGGGCGCGCGCATGCGGGGGCCATCGCGAGGGCCTCATCAGAGCAG 120  
Db 61 TGGGCTTCGTTGGGCGCGCGCGCATGCGGGGGCCATCGCGAGGGCCTCATCAGAGCAG 120  
Qy 121 GAAAGTGGAGGCTCAGCACAATCTGCGCAGTGCACCAACAGACAGGAACCTATGTCACT 180  
Db 121 GAAAGTGGAGGCTCAGCACAATCTGCGCAGTGCACCAACAGACAGGAACCTATGTCACT 180  
Qy 181 TTCAAGCTCTGGGTTGCGGACCAACGACATCCCAACCAAGAGGTTGCTGCAGAGTGCCTGC 240  
Db 181 TTCAAGCTCTGGGTTGCGGACCAACGACATCCCAACCAAGAGGTTGCTGCAGAGTGCCTGC 240  
Qy 241 TCCTCATCTTTGGCACCACCTCATGCTGCCAGCTGCTCCGCGAGAGTGGCTCCCG 300  
Db 241 TCCTCATCTTTGGCACCACCTCATGCTGCCAGCTGCTCCGCGAGAGTGGCTCCCG 300  
Qy 301 TGGTTCACCATGAAACACATCTTGGTGTCCGTGGGCTGCTCTGAGCACCCCTGG 360  
Db 301 TGGTTCACCATGAAACACATCTTGGTGTCCGTGGGCTGCTCTGAGCACCCCTGG 360  
Qy 361 AGGAGCTGTGCCCCCAACACACAGGGTGTCTGGGGTCTTGGCCAACTGCCCTGTGTGG 420  
Db 361 AGGAGCTGTGCCCCCAACACACAGGGTGTCTGGGGTCTTGGCCAACTGCCCTGTGTGG 420  
Qy 421 TCCAGGAAGGGCCATAGTATGCGCGGGGGCCACGTGGGGACCGAGACCAACG 480  
Db 421 TCCAGGAAGGGCCATAGTATGCGCGGGGGCCACGTGGGGACCGAGACCAACG 480  
Qy 481 TCCTGCAGCATCTGCTGGAGGCCCTGTGGGCGGTGTGAGGAGGTGCTGAAGCCTACGTGG 540  
Db 481 TCCTGCAGCATCTGCTGGAGGCCCTGTGGGCGGTGTGAGGAGGTGCTGAAGCCTACGTGG 540

QY 541 ACATCCACACTGGCCTCAGTGGCAGTGGCGTGGCCTTCGTGTGTGTCATCTCCGAGGCC 600  
Db 541 ACATCCACACTGGCCTCAGTGGCAGTGGCGTGGCCTTCGTGTGTGTCATCTCCGAGGCC 600  
QY 601 TGGCTGAAGGAGCCCTCAAGATGGCATGCCAGCAGCCTGGCCACCGCATCGCTGCC 660  
Db 601 TGGCTGAAGGAGCCCTCAAGATGGCATGCCAGCAGCCTGGCCACCGCATCGCTGCC 660  
QY 661 AGACCCCTGCTGGGAGCGCCAGATGCTGTCAGAGGAGGCAACACCCAGCCAGCTGC 720  
Db 661 AGACCCCTGCTGGGAGCGCCAGATGCTGTCAGAGGAGGCAACACCCAGCCAGCTGC 720  
QY 721 GCTCAGAGTGTGCACCCCGGGTGGCACCAACCATATATGGACTCCACCCCTGGAGCAGG 780  
Db 721 GCTCAGAGTGTGCACCCCGGGTGGCACCAACCATATATGGACTCCACCCCTGGAGCAGG 780  
QY 781 GCGGCTCGGAGCAGCCACCATGAGCGCGTGGAGGCTGCCACCTGCCGGGCCAAGGAGC 840  
Db 781 GCGGCTCGGAGCAGCCACCATGAGCGCGTGGAGGCTGCCACCTGCCGGGCCAAGGAGC 840  
QY 841 TCAGCAGAAAGTAGGCTGGGCTCTGGCCATCCTTTCCCTGCTGTGCTGCTGCTCTCC 900  
Db 841 TCAGCAGAAAGTAGGCTGGGCTCTGGCCATCCTTTCCCTGCTGTGCTGCTGCTCTCC 900  
QY 901 CTGTCTCCCTTCCCTGAGGACTGCGGCTCCCTCCCTGCTGCATGAGGTTCTCTACTGC 960  
Db 901 CTGTCTCCCTTCCCTGAGGACTGCGGCTCCCTCCCTGCTGCATGAGGTTCTCTACTGC 960  
QY 961 TCCTTCTCCCTTCCGACAGGAAATGACAGGGGAGGACTTGGGAGGTTCCAGCAGCGG 1020  
Db 961 TCCTTCTCCCTTCCGACAGGAAATGACAGGGGAGGACTTGGGAGGTTCCAGCAGCGG 1020  
QY 1021 GGGAGCCCCGACAGTGGGACACTCCCTCCCTCCCGAGTGAGCAGAGGACCCGTGGTGG 1080  
Db 1021 GGGAGCCCCGACAGTGGGACACTCCCTCCCTCCCGAGTGAGCAGAGGACCCGTGGTGG 1080  
QY 1081 TGGCTCTCCCTTGGCTGCAGTGGAGCCACCTTGTGCAACATTTGTTCTGAGGGGCCCA 1140  
Db 1081 TGGCTCTCCCTTGGCTGCAGTGGAGCCACCTTGTGCAACATTTGTTCTGAGGGGCCCA 1140  
QY 1141 AGAGATGGCTTGGTTCATTTGCCCGCATGTTGGGCGAGTGGTGGGCGCATGAACAG 1200  
Db 1141 AGAGATGGCTTGGTTCATTTGCCCGCATGTTGGGCGAGTGGTGGGCGCATGAACAG 1200  
QY 1201 AACTTACGGTAACAGGCACGCTGGCCCAATGCTGCTGGAGCTGGAGCTTGCCTTTG 1260  
Db 1201 AACTTACGGTAACAGGCACGCTGGCCCAATGCTGCTGGAGCTGGAGCTTGCCTTTG 1260  
QY 1261 GCTTTCCAAGTGGGCTCGTGCAGCTACAGCCAGGCGCGCTGCCTCATCTCAGCTCTAGG 1320  
Db 1261 GCTTTCCAAGTGGGCTCGTGCAGCTACAGCCAGGCGCGCTGCCTCATCTCAGCTCTAGG 1320  
QY 1321 GGCAGGACATATGGGT 1338  
Db 1321 GGCAGGACATATGGGT 1338  
RESULT 2  
AAH18288  
ID AAH18288 standard; cDNA; 2331 BP.  
XX  
XX AAH18288;  
XX AC  
XX DT 26-JUN-2001 (first entry)  
XX DE  
XX Human cDNA sequence SEQ ID NO:18268.  
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.  
XX Homo sapiens.  
XX XX  
PN EP1074617-A2.

XX PD. 07-FEB-2001.  
XX PF. 28-JUL-2000; 2000EP-0116126.  
XX PR. 29-JUL-1999; 99JP-0248036.  
XX PR. 27-AUG-1999; 99JP-0300253.  
XX PR. 11-JAN-2000; 2000JP-0118776.  
XX PR. 02-MAY-2000; 2000JP-0183767.  
XX PR. 09-JUN-2000; 2000JP-0241899.  
XX PA. (HELI-) HELIX RES INST.  
XX PI. Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
XX WPI; 2001-318749/34.  
XX PT. Primer sets for synthesizing polynucleotides, particularly the 5602  
PT full-length cDNAs defined in the specification, and for the detection  
PT and/or diagnosis of the abnormality of the proteins encoded by the  
PT full-length cDNAs -  
XX PS. Claim 8; SEQ ID 18268; 2537pp + CD ROM; English.  
XX CC. The present invention describes primer sets for synthesizing 5602  
CC full-length cDNAs defined in the specification. Where a primer set  
CC comprises: (a) an oligo-dr primer and an oligonucleotide complementary  
CC to the complementary strand of a polynucleotide which comprises one of  
CC the 5602 nucleotide sequences defined in the specification, where the  
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
CC of an oligonucleotide comprising a sequence complementary to the  
CC complementary strand of a polynucleotide which comprises a 5'-end  
CC sequence and an oligonucleotide comprising a sequence complementary to a  
CC polynucleotide which comprises a 3'-end sequence, where the  
CC oligonucleotide comprises at least 15 nucleotides and the combination of  
CC the 5'-end sequence/3'-end sequence is selected from those defined in  
CC the specification. The primer sets can be used in antisense therapy and  
CC in gene therapy. The primers are useful for synthesizing polynucleotides,  
CC particularly full-length cDNAs. The primers are also useful for the  
CC detection and/or diagnosis of the abnormality of the proteins encoded by  
CC the full-length cDNAs. The primers allow obtaining of the full-length  
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to  
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632  
CC represent oligonucleotides, all of which are used in the exemplification  
CC of the present invention.  
XX SQ. Sequence 2331 BP; 398 A; 753 C; 748 G; 432 T; 0 other;  
Query Match 97.7%; Score 1307; DB 22; Length 2331;  
Best Local Similarity 99.2%; Pred. No. 3.3e-282;  
Matches 1313; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
Qy 15 GTCGAGGCAACAAGATGGCAGCTGGCGAGCCGCTCCCGCGCGCTGGCTTCGTGGGC 74  
Db 12 GTCGAGGCAACAAGATGGCAGCTGGCGAGCCGCTCCCGCGCGCTGGCTTCGTGGGC 71  
Qy 75 GCGGGCGCATGCGCGGGGCGATCGCGCAGGGCCTCATCAGACGAGGAAAGTGAAGCT 134  
Db 72 GCGGGCGCATGCGCGGGGCGCATCGCGCAGGGCCTCATCAGACGAGGAAAGTGAAGCT 131  
Qy 135 CAGCACAATAGTGGCCAGTGCACCAACAGACAGAACCTATGTCACCTTCAAGCTCTGGGT 194  
Db 132 CAGCACAATAGTGGCCAGTGCACCAACAGACAGAACCTATGTCACCTTCAAGCTCTGGGT 191  
Qy 195 TGCGGGACCAACGACCTCCCAACAGGAGGTGCTGCAGAGCTGCTGCTCATCTTTGTC 254  
Db 192 TGCGGGACCAACGACCTCCCAACAGGAGGTGCTGCAGAGCTGCTGCTCATCTTTGTC 251  
Qy 255 ACCAAGCCTCATGTGCTGCCAGCTGTCTTCCTGGCAGAGGTGGCTGCTGCTACCACTGAA 314  
Db 252 ACCAAGCCTCATGTGCTGCCAGCTGTCTTCCTGGCAGAGGTGGCTGCTGCTACCACTGAA 311

Qy 315 CACATCTTGGTGTCCGTGGCTGCTGGGGTGTCTCTGAGCACCCTGGAGGAGCTGCTGCC 374  
Db 312 CACATCTTGGTGTCCGTGGCTGCTGGGGTGTCTCTGAGCACCCTGGAGGAGCTGCTGCC 371  
Qy 375 CCAAAACACAGGGTGTCTGCCGGTCTTGCACCACTGCCCTGTGTGGTCCAGGAAGGGGCC 434  
Db 372 CCAAAACACAGGGTGTCTGCCGGTCTTGCACCACTGCCCTGTGTGGTCCAGGAAGGGGCC 431  
Qy 435 ATAGTGATGGCGCGGGGCGCCACGTGGGAGCAGCAGACCAAGCTCTCTGACGATCTG 494  
Db 432 ATAGTGATGGCGCGGGGCGCCACGTGGGAGCAGCAGACCAAGCTCTCTGACGATCTG 491  
Qy 495 CTGAGAGCCTGTGGCGGCTGTGAGAGGTGCTTGAAGCTACGTCGACATCCACACATGGC 554  
Db 492 CTGAGAGCCTGTGGCGGCTGTGAGAGGTGCTTGAAGCTACGTCGACATCCACACATGGC 551  
Qy 555 CTCAGTGGCAGTGGCGGCTTCTGTGTGTCATTCTCCGAGGCCCTGGCTGAAGAGGCC 614  
Db 552 CTCAGTGGCAGTGGCGGCTTCTGTGTGTCATTCTCCGAGGCCCTGGCTGAAGAGGCC 611  
Qy 615 GTCAGATGGCAGTGGCGGCTTCTGTGTGTCATTCTCCGAGGCCCTGGCTGAAGAGGCC 674  
Db 612 GTCAGATGGCAGTGGCGGCTTCTGTGTGTCATTCTCCGAGGCCCTGGCTGAAGAGGCC 671  
Qy 675 ACGCCAAGATGCTGTCAGAGGCGCAACACCCAGCCAGCTGCGCTCAGAGCTGTGC 734  
Db 672 ACGCCAAGATGCTGTCAGAGGCGCAACACCCAGCCAGCTGCGCTCAGAGCTGTGC 731  
Qy 735 ACCCGGGTGGCAGCACCATCATATGGACTCCACGCCCTGGAGCAGGGCGGCTCGGAGCA 794  
Db 732 ACCCGGGTGGCAGCACCATCATATGGACTCCACGCCCTGGAGCAGGGCGGCTCGGAGCA 791  
Qy 795 GCCACATGAGCGCGGTGGAGGCTGCCACCTGCGGGGCCAAGAGCTCAGCAGAAAGTAG 854  
Db 792 GCCACATGAGCGCGGTGGAGGCTGCCACCTGCGGGGCCAAGAGCTCAGCAGAAAGTAG 851  
Qy 855 GCTGGGCTCTGGCCATCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 914  
Db 852 GCTGGGCTCTGGCCATCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 911  
Qy 915 CTGAGGACTCGGGTCTCCCTCCTGTCATGAGGGTCTCTTCTGCTGCTGCTGCTGCTGCT 974  
Db 912 CTGAGGACTCGGGTCTCCCTCCTGTCATGAGGGTCTCTTCTGCTGCTGCTGCTGCTGCT 971  
Qy 975 CACAGGAAATGCGAGGGGCGAGGCTTGGAGGTTCCAGCAGCGGGGAGGCCCGGAGCA 1034  
Db 972 CACAGGAAATGCGAGGGGCGAGGCTTGGAGGTTCCAGCAGCGGGGAGGCCCGGAGCA 1031  
Qy 1035 GTGGGGACACTCTCCCTCCCGAGTGAAGAGCAGCCGTGGTGGTGGCTTCTGCCCTT 1094  
Db 1032 GTGGGGACACTCTCCCTCCCGAGTGAAGAGCAGCCGTGGTGGTGGCTTCTGCCCTT 1091  
Qy 1095 GCTCAGTGAAGCCACCTTGTGCAACATTTGGTCTGAGGGGCCCCAAGAGATGGGCTTT 1154  
Db 1092 GCTCAGTGAAGCCACCTTGTGCAACATTTGGTCTGAGGGGCCCCAAGAGATGGGCTTT 1151  
Qy 1155 GGTCAATTTCCCGCATGTTGGCAGTTGGTGGCCATGAGGCTTGAAGCTTACGATACA 1214  
Db 1152 GGTCAATTTCCCGCATGTTGGCAGTTGGTGGCCATGAGGCTTGAAGCTTACGATACA 1211  
Qy 1215 GGCAGGCTGGCCCCAATGCTGCTGGAGCTTGGAGCTTGGCTTTGGCTTTTCCCAAGTGG 1274  
Db 1212 GGCAGGCTGGCCCCAATGCTGCTGGAGCTTGGAGCTTGGCTTTGGCTTTCCAGGTGG 1271  
Qy 1275 CTCGTGAGCTACAGCCAGGCGGCTGCTCATCTTCTAGCTCTAGGGGCGCAGGATATG 1334  
Db 1272 TCCGTGAGCTACAGCCAGGCGGCTGCTCATCTCTAGCTCTAGGGGCGCAGGATATG 1331  
Qy 1335 GGG 1337  
Db 1332 GGG 1334

RESULT 3

ABA00759  
ID ABA00759 standard; DNA; 2331 BP.

XX AC ABA00759;  
XX AC

DT 18-MAR-2003 (first entry)  
XX DE

DE P5CR related nucleic acid #2.  
XX KW

KW Pyroline 5 carboxylate reductase; P5CR; NADPH; cancer; breast; colon;  
KW Pyroline 5 carboxylate; P5C; proline; p53; kidney; lung; ovary; ss.  
XX OS

XX Homo sapiens.  
XX PN

PN W0200299043-A2.  
XX PD

PD 12-DEC-2002.  
XX XX

XX 03-JUN-2002; 2002W0-US17319.  
PF XX

XX 05-JUN-2001; 2001US-296080P.  
PR PR

PR 10-OCT-2001; 2001US-328509P.  
XX XX

XX (EXEL-) EXELIS INC.  
PA XX

XX Friedman L, Plowman GD, Belvin M, Francis-Lang H, Li D, Funke RP;  
PI Engst S;  
XX XX

XX WPI; 2003-140606/13.  
DR DR

XX Identifying p53 pathway modulating agents with p5CR genes, useful for  
PT the diagnosis and treatment of disorders associated with defects in the  
PT p53 pathway, such as cancer of the breast, colon, kidneys, lung and  
PT ovary  
XX XX

XX Disclosure; Page 39-40; 53pp; English.  
XX XX

CC The sequences given in ABA00758-66 represent pyrroline 5 carboxylate  
CC reductase (P5CR) related nucleic acid sequences. P5CR catalyzes the  
CC NAD(P)H-dependent conversion of pyrroline 5 carboxylate (P5C) to  
CC proline. These sequences may be used in the method of the invention  
CC for identifying a candidate p53 pathway modulating agent. The method  
CC comprises providing an assay system comprising a purified P5CR  
CC polypeptide or nucleic acid, or a functionally active fragment or  
CC system provides a reference activity, and detecting a test agent-biased  
CC activity of the assay system. The method of the present invention is  
CC useful for the diagnosis and treatment of disorders associated with  
CC defects in the p53 pathway, such as cancer of the breast, colon,  
CC kidneys, lung and ovary.  
XX XX

SQ Sequence 2331 BP; 398 A; 753 C; 748 G; 432 T; 0 other;

Query Match 97.7%; Score 1307; DB 25; Length 2331;  
Best Local Similarity 99.2%; Pred. No. 3.3e-282;  
Matches 1313; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 15 GTCCGAGGCAACAAGATGCGAGCTCGGAGCGCTCTCCGCGCGGTGGCTCGTGGGC 74  
DB 12 GTCCGAGGCAACAAGATGCGAGCTCGGAGCGCTCTCCGCGCGGTGGCTCGTGGGC 71

QY 75 GCGGGCCGATGCGGGGGGCCATCGCGAGGCGCTCATCAGACGAGGAAAAGTGAAGCT 134  
DB 72 GCGGGCCGATGCGGGGGGCCATCGCGAGGCGCTCATCAGACGAGGAAAAGTGAAGCT 131

QY 135 CAGCACATCTGGCCAGTGCACCAACACAGAGAACCTATGTACTTTCAAGCTGTGGGT 194  
DB 132 CAGCACATCTGGCCAGTGCACCAACACAGAGAACCTATGTACTTTCAAGCTGTGGGT 191

QY 195 TCCCGGACACGCACTCCAACACAGGAGGTGTGACAGAGCTGCCTGCTCATCTTGGC 254  
DB 192 TCCCGGACACGCACTCCAACACAGGAGGTGTGACAGAGCTGCCTGCTCATCTTGGC 251

DB 192 TCCCGGACACGCACTCCAACACAGGAGGTGTGACAGAGCTGCCTGCTCATCTTGGC 251  
QY 255 ACCAAGCCTCATGTGCTGCCAGCTGTCTGTGCAGAGTGGCTCTGTGGTCAACACAGTAA 314  
DB 252 ACCAAGCCTCATGTGCTGCCAGCTGTCTGTGCAGAGTGGCTCTGTGGTCAACACAGTAA 311  
QY 315 CACATCTTGTGTCTCCGTGGCTGTCTGTAGACACCCCTGTGAGGAGCTGTCTGCC 374  
DB 312 CACATCTTGTGTCTCCGTGGCTGTCTGTAGACACCCCTGTGAGGAGCTGTCTGCC 371  
QY 375 CCAACACACGGGTGTCTGCCAACCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 434  
DB 372 CCAACACACGGGTGTCTGCCAACCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 431  
QY 435 ATAGTGTATGGCGGGCGCCACGTCGTGGGAGCAGACACAGCTCTGTGACAGCTGTG 494  
DB 432 ATAGTGTATGGCGGGCGCCACGTCGTGGGAGCAGCAGACCAAGCTCTGTGACAGCTGTG 491  
QY 495 CTGAGGCGCTGTGGCGGTGTGTGAGGAGTGTCTGAAGCCTACGTCGACATCCACACTGGC 554  
DB 492 CTGAGGCGCTGTGGCGGTGTGTGAGGAGTGTCTGAAGCCTACGTCGACATCCACACTGGC 551  
QY 555 CTCAGTGGCAGTGGCGCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 614  
DB 552 CTCAGTGGCAGTGGCGCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 611  
QY 615 GTCAAGATGGCAGTGGCGCGCCACGTCGTGGGAGCAGCAGCTGTGACAGCTGTGTGTGTGT 674  
DB 612 GTCAAGATGGCAGTGGCGCGCCACGTCGTGGGAGCAGCAGCTGTGACAGCTGTGTGTGTGT 671  
QY 675 ACGGCCAAGATGT 734  
DB 672 ACGGCCAAGATGT 731  
QY 735 ACCCGGGTGGCAGCAGCTGT 794  
DB 732 ACCCGGGTGGCAGCAGCTGT 791  
QY 795 GCCACCATGAGCGCGGTGGAGGT 854  
DB 792 GCCACCATGAGCGCGGTGGAGGT 851  
QY 855 GCTGGGCTGTGGCGATCTTTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 914  
DB 852 GCTGGGCTGTGGCGATCTTTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 911  
QY 915 CTGAGGACTGGGCT 974  
DB 912 CTGAGGACTGGGCT 971  
QY 975 CACAGGAAATGCGAGGCGAGGACTTGGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1034  
DB 972 CACAGGAAATGCGAGGCGAGGACTTGGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1031  
QY 1035 GTGGGAGACT 1094  
DB 1032 GTGGGAGACT 1091  
QY 1095 GCTGCACTGAGCCCACTTCT 1154  
DB 1092 GCTGCACTGAGCCCACTTCT 1151  
QY 1155 GGTCACTTGGCCGATGGTGGGCACTTGGTGTGAGGCGCATGAACAGAACTTACGGTAAACA 1214  
DB 1152 GGTCACTTGGCCGATGGTGGGCACTTGGTGTGAGGCGCATGAACAGAACTTACGGTAAACA 1211  
QY 1215 GGCAGGCTGGCCCAATGT 1274  
DB 1212 GGCAGGCTGGCCCAATGT 1271  
QY 1275 CTGCTGACGTACAGCCAGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1334  
DB 1272 TCCGTGACGTACAGCCAGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1331

QY 1335 GGG 1337  
Db 1332 GGG 1334

## RESULT 4

AAH77597

ID AAH77597 standard; cDNA; 2496 BP.

XX AC AAH77597;

XX DT 22-OCT-2001 (first entry)

XX DE Human dihydropyrrrole-5-carboxylate reductase 30 cDNA.

XX KW Human; dihydropyrrrole-5-carboxylate reductase 30; cancer; cytostatic;

XX KW human immunodeficiency virus; HIV; infection; immunological disease;

XX KW inflammatory disease; ss.

XX OS Homo sapiens.

XX PN CN1298002-A.

XX PD 06-JUN-2001.

XX PF 24-NOV-1999; 99CN-0124090.

XX PR 24-NOV-1999; 99CN-0124090.

XX PA (SHAN-) SHANGHAI BORONG GENE DEV CO LTD.

XX PI Mao Y, Xie Y;

XX DR WPI; 2001-489680/54.

XX DR P-PSDB; AAG66956.

XX PT Human dihydropyrrrole-5-carboxylate reductase 30 as one new kind of

XX PS polypeptide and polynucleotides encoding this polypeptide -

XX CS Claim 6; Page 19-20 (disclosure); 26pp; Chinese.

XX CC The invention relates to a novel polypeptide, human

XX CC dihydropyrrrole-5-carboxylate reductase 30, polynucleotides encoding

XX CC this polypeptide and a DNA recombination process to produce the

XX CC polypeptide. The polypeptide is useful for treating various diseases,

XX CC such as malignant tumours, nosohaemia, HIV infection, immunological

XX CC diseases and inflammatory diseases. The invention also provides an

XX CC antibody against the polypeptide. The present sequence encodes the

XX CC polypeptide of the invention.

XX SQ Sequence 2496 BP; 451 A; 793 C; 789 G; 463 T; 0 other;

Query Match 97.3%; Score 1302.2; DB 22; Length 2496;  
Best Local Similarity 99.0%; Pred. No. 3.9e-281;  
Matches 1310; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 15 GTCCGAGGACACAGATGCGAGTGGGCGGCTCCCGCGCGGTGGGCTTCGTGGGC 74  
|||||  
Db 14 GTCCGAGGACACAGATGCGAGTGGGCGGCTCCCGCGCGGTGGGCTTCGTGGGC 73  
|||||  
QY 75 GCGGGCGGATGCGGGGGCCATCGCGAGGGCCCTCATCAGAGCAGGAAAAGTGAAGCT 134  
|||||  
Db 74 GCGGGCGGATGCGGGGGCCATCGCGAGGGCCCTCATCAGAGCAGGAAAAGTGAAGCT 133  
|||||  
QY 135 CAGCACATAGTGGCCAGTGCACCAACAGACAGAACTATGTCACCTTTCAAGCTCTGGGT 194  
|||||  
Db 134 CAGCACATAGTGGCCAGTGCACCAACAGACAGAACTATGTCACCTTTCAAGCTCTGGGT 193  
|||||  
QY 195 TGGCGGACACGACACTCCAAACAGAGGTGCTGCAGAGTGCCTGCTGTCATCTTTGCC 254  
|||||  
Db 194 TGGCGGACACGACACTCCAAACAGAGGTGCTGCAGAGTGCCTGCTGTCATCTTTGCC 253  
|||||

|            |   |          |    |      |  |
|------------|---|----------|----|------|--|
| Db         | 1334  | GGG 1336 | QY | 27   | AAGATGCGAGCTCGGAGCCGCTCTCCGCGCGCTGGGCTTCGTGGCGCGGCGCCGATG 86                 |
|            |   |          | Db | 2615 | AAGATGCGAGCTCGGAGCCGCTCTCCGCGCGCTGGGCTTCGTGGCGCGGCGCCGATG 2556               |
| RESULT 5   |   |          | QY | 87   | GCGGGGCGATCGCGCAGGCGCTCATCAGAGCAGGAAAGTGAAGCTCAGCACACTACTG 146               |
| AAH14370/c |   |          | Db | 2555 | GCGGGGCGATCGCGCAGGCGCTCATCAGAGCAGGAAAGTGAAGCTCAGCACACTACTG 2496              |
| XX         | AAH14370;   |          | QY | 147  | GCCAGTGCACCAACAGACAGAGAACCTATGTCACCTTTCACCTTTCAGCTCTGGGTTGCCGGACACG 206      |
| XX         | AC  |          | Db | 2495 | GCCAGTGCACCAACAGACAGAGAACCTATGTCACCTTTCACCTTTCAGCTCTGGGTTGCCGGACACG 2436     |
| XX         | DT  |          | QY | 207  | CACCTCCAAACAGAGAGTGCCTGCAGAGCTGCTGCTCATCTTTCACCTTTCAGCTCTGGGTTGCCGGACACG 266 |
| XX         | 26-JUN-2001 (first entry)   |          | Db | 2435 | CACCTCCAAACAGAGAGTGCCTGCAGAGCTGCTGCTCATCTTTCAGCTCTGGGTTGCCGGACACG 2376       |
| XX         | Human cDNA sequence SEQ ID NO:11778.                                      |          | QY | 267  | GTGCTGCCAGTGTCTGGCAGAGGTGGCTCTGTGGTTCACCTACCTGACACACTTCTTGGTG 326            |
| XX         | Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss. |          | Db | 2375 | GTGCTGCCAGTGTCTGGCAGAGGTGGCTCTGTGGTTCACCTACCTGACACACTTCTTGGTG 2316           |
| XX         | Homo sapiens.   |          | QY | 327  | TCCGTGCTGCTGGGGTGTCTCTGAGCACCTTGAGGAGCTGCTGCCCCCAACACACG 386                 |
| XX         | EP1074617-A2.   |          | Db | 2315 | TCCGTGCTGCTGGGGTGTCTCTGAGCACCTTGAGGAGCTGCTGCCCCCAACACACG 2256                |
| XX         | 07-FEB-2001.  |          | QY | 387  | GTGCTGCGGGTCTTCCCAACCTGCTGTGTGTCCAGGAAGGGCCATAGTAGGGG 446                    |
| XX         | 28-JUL-2000; 2000EP-0116126.  |          | Db | 2255 | GTGCTGCGGGTCTTCCCAACCTGCTGTGTGTCCAGGAAGGGCCATAGTAGGGG 2196                   |
| XX         | 29-JUL-1999; 99JP-0248036.  |          | QY | 447  | CGGGCCGCGACGTGGGGAGCAGGAGCAACGCTCTGCAGCATCTGCTGGAGGCGCTGT 506                |
| XX         | 27-AUG-1999; 99JP-0300253.  |          | Db | 2195 | CGGGCCGCGACGTGGGGAGCAGGAGCAACGCTCTGCAGCATCTGCTGGAGGCGCTGT 2136               |
| XX         | 11-JAN-2000; 2000JP-0118776.  |          | QY | 507  | GGGCGGTGTGAGGAGGTGCGTGAAGCTACGTCGACATCCACACTGCGCTCAGTGCAGT 566               |
| XX         | 02-MAY-2000; 2000JP-0183767.  |          | Db | 2135 | GGGCGGTGTGAGGAGGTGCGTGAAGCTACGTCGACATCCACACTGCGCTCAGTGCAGT 2076              |
| XX         | 09-JUN-2000; 2000JP-0241899.  |          | QY | 567  | GGCGTGCCCTTCGTGTGTGCAATTCGAGGCGCTGGCTGAAGAGCGCTCAAGATGGG 626                 |
| XX         | (HELI-) HELIX RES INST.   |          | Db | 2075 | GGCGTGCCCTTCGTGTGTGCAATTCGAGGCGCTGGCTGAAGAGCGCTCAAGATGGG 2016                |
| XX         | Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;             |          | QY | 627  | ATGCCAGCAGCCTGGCCACCGCATCGCTGCCACACACCTGCTGGGGAGCGGCAAGATG 686               |
| XX         | Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;                      |          | Db | 2015 | ATGCCAGCAGCCTGGCCACCGCATCGCTGCCACACACCTGCTGGGGAGCGGCAAGATG 1956              |
| XX         | WPI; 2001-318749/34.  |          | QY | 687  | CTGCTGCAGCAGGCGCAACACCCAGCCAGCTGCGCTCAGACGTGTGCACCCCGGTGGC 746               |
| XX         | Primer sets for synthesizing polynucleotides, particularly the 5602       |          | Db | 1955 | CTGCTGCAGCAGGCGCAACACCCAGCCAGCTGCGCTCAGACGTGTGCACCCCGGTGGC 1896              |
| XX         | full-length cDNAs defined in the specification, and for the detection     |          | QY | 747  | ACCACCATCTATGAGCTTCCACGCCCTGGAGCAGGCGGGCTGGAGCAGCAGCATGAGC 806               |
| XX         | and/or diagnosis of the abnormality of the proteins encoded by the        |          | Db | 1895 | ACCACCATCTATGAGCTTCCACGCCCTGGAGCAGGCGGGCTGGAGCAGCAGCATGAGC 1836              |
| XX         | full-length cDNAs -   |          | QY | 807  | GCGTGAGGCTGCCACCTGCGCGGCGCAAGGAGCTCAGCAGAAAGTAGGCTGGCTCG 866                 |
| XX         | Claim 8; SEQ ID 11778; 2537pp + CD ROM; English.                          |          | Db | 1835 | GCGTGAGGCTGCCACCTGCGCGGCGCAAGGAGCTCAGCAGAAAGTAGGCTGGCTCG 1776                |
| XX         | The present invention describes primer sets for synthesizing 5602         |          | QY | 867  | CCATCTCTTCTGCTGCTGCGCCCTCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 926               |
| XX         | full-length cDNAs defined in the specification. Where a primer set        |          | Db | 1775 | CCATCTCTTCTGCTGCTGCGCCCTCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1716              |
| XX         | comprises: (a) an oligo-dr primer and an oligonucleotide complementary    |          | QY | 927  | GCTCCCTCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 986              |
| XX         | to the complementary strand of a polynucleotide which comprises one of    |          | Db | 1715 | GCTCCCTCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1656             |
| XX         | the 5602 nucleotide sequences defined in the specification, where the     |          | QY | 987  | CAGGGGCGAGGCTTGGAGGTTCCAGAGCGGGGAGCCCGGAGCTGGGAGGAGCTTC 1046                 |
| XX         | oligonucleotide comprises at least 15 nucleotides; or (b) a combination   |          | Db | 1655 | CAGGGGCGAGGCTTGGAGGTTCCAGAGCGGGGAGCCCGGAGCTGGGAGGAGCTTC 1596                 |
| XX         | of an oligonucleotide comprising a sequence complementary to the          |          | QY | 1047 | CTCCCTCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1106              |
| XX         | complementary strand of a polynucleotide which comprises a 5'-end         |          | Db | 1595 | CTCCCTCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1536              |
| XX         | sequence and an oligonucleotide comprising a sequence complementary to a  |          | QY | 1107 | CCACTTGTGTCAACATTTGTTCTGAGGGGCGCCCAAGAGATGGGCTCTTGGTCATTTGGCC 1166           |
| XX         | polynucleotide which comprises a 3'-end sequence, where the               |          |    |      |  |
| XX         | oligonucleotide comprises at least 15 nucleotides, and the combination of |          |    |      |  |
| XX         | the 5'-end sequence/3'-end sequence is selected from those defined in     |          |    |      |  |
| XX         | the specification. The primer sets can be used in antisense therapy and   |          |    |      |  |
| XX         | in gene therapy. The primers are useful for synthesizing polynucleotides, |          |    |      |  |
| XX         | particularly full-length cDNAs. The primers are also useful for the       |          |    |      |  |
| XX         | detection and/or diagnosis of the abnormality of the proteins encoded by  |          |    |      |  |
| XX         | the full-length cDNAs. The primers allow obtaining of the full-length     |          |    |      |  |
| XX         | cDNAs easily without any specialised methods. AAH03166 to AAH13628 and    |          |    |      |  |
| XX         | AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to          |          |    |      |  |
| XX         | AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632   |          |    |      |  |
| XX         | represent oligonucleotides, all of which are used in the exemplification  |          |    |      |  |
| XX         | of the present invention.   |          |    |      |  |
| XX         | Sequence 2615 BP; 492 A; 831 C; 818 G; 474 T; 0 other;                    |          |    |      |  |
| XX         | Query Match   |          |    |      |  |
| XX         | Best Local Similarity 96.5%; Score 1291.8; DB 22; Length 2615;            |          |    |      |  |
| XX         | Matches 1299; Conservative 0; Mismatches 12; Indels 0; Gaps 0;            |          |    |      |  |



Db 1535 CCACCTTGCCTCCCAACATTTGGTTCTGAGGGGCCCAAGAGATGGCTCTTGGTCAATTTGCC 1476  
QY 1167 GCATGGTTGGGCACTTGGTTGAGGCCATGAACAGAACTTACGGTTAACAGCAGCGGCTGGC 1226  
Db 1475 GCATGGTTGGGCACTTGGTTGAGGCCATGAACAGAACTTACGGTTAACAGCAGCGGCTGGC 1416  
QY 1227 CCAATGCTGGTCTGGAGCTGGAGCTTGCCTTTTGGCTTTCCAAAGTGGGCTTCGTGCAGCTA 1286  
Db 1415 CCAATGCTGGTCTGGAGCTGGAGCTTGCCTTTTGGCTTTCCAGTGGCTCCGTGCAGCTA 1356  
QY 1287 CAGCCAGCGGCTGCCTCATCTCAGCTCTAGGGGGCAGCAGCATATFGGG 1337  
Db 1355 CAGCCAGCGGCTGCCTCATCTCAGCTCTAGGGGGCAGCAGCATATFGGG 1305

RESULT 6  
ABA00764  
ID ABA00764 standard; DNA: 1178 BP.  
XX ABA00764;  
AC  
XX  
XX  
DT 18-MAR-2003 (first entry)  
XX  
XX  
DE P5CR related nucleic acid #7.  
XX  
KW Pyroline 5 carboxylate reductase; P5CR; NADPH; cancer; breast; colon;  
KW pyroline 5 carboxylate; P5C; proline; p53; kidney; lung; ovary; ss.  
XX  
OS Homo sapiens.  
XX  
XX WO200299043-A2.  
XX  
XX 12-DEC-2002.  
XX  
XX 03-JUN-2002; 2002WO-US17319.  
XX  
XX 05-JUN-2001; 2001US-296080P.  
PR 10-OCT-2001; 2001US-328509P.  
XX  
XX (EXEL-) EXELIS INC.  
XX

PI Friedman L, Plowman GD, Belvin M, Francis-Lang H, Li D, Funke RP;  
PI Engst S;  
XX  
XX WPI; 2003-140606/13.  
DR  
XX  
XX Identifying p53 pathway modulating agents with p5CR genes, useful for  
PT the diagnosis and treatment of disorders associated with defects in the  
PT p53 pathway, such as cancer of the breast, colon, kidneys, lung and  
PT ovary -  
XX  
XX Disclosure; Page 44-45; 53pp; English.

XX The sequences given in ABA00758-66 represent pyroline 5 carboxylate  
CC reductase (P5CR) related nucleic acid sequences. P5CR catalyses the  
CC NAD(P)H-dependent conversion of pyroline 5 carboxylate (P5C) to  
CC proline. These sequences may be used in the method of the invention  
CC for identifying a candidate p53 pathway modulating agent. The method  
CC comprises providing an assay system comprising a purified P5CR  
CC polypeptide or nucleic acid, or a functionally active fragment or  
CC derivative, contacting the assay system with a test agent, where the  
CC system provides a reference activity, and detecting a test agent-biased  
CC activity of the assay system. The method of the present invention is  
CC useful for the diagnosis and treatment of disorders associated with  
CC defects in the p53 pathway, such as cancer of the breast, colon,  
CC kidneys, lung and ovary.

SQ Sequence 1178 BP; 211 A; 379 C; 383 G; 205 T; 0 other;

Query Match 84.3%; Score 1127.8; DB 25; Length 1178;  
Best Local Similarity 99.8%; Pred. No. 2.8e-242;  
Matches 1129; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 15 GTCCGAGGCAACAAGATGCAGTCGGAGCCGCTCTCCCGCGCGGTGGCTTCGTGGGC 74  
Db 32 GTCCGAGGCAACAAGATGCAGTCGGAGCCGCTCTCCCGCGCGGTGGCTTCGTGGGC 91  
QY 75 GCGGGCCGATGGCGGGGCCATCGCGCAGGGCCCTCATCAGAGCAGGAAAGTGAAGCT 134  
Db 92 GCGGGCCGATGGCGGGGCCATCGCGCAGGGCCCTCATCAGAGCAGGAAAGTGAAGCT 151  
QY 135 CAGCACATCTGGCCAGTCCACACAGACAGAACCTATGTACTTTCAAGCTCTGGGT 194  
Db 152 CAGCACATCTGGCCAGTCCACACAGACAGAACCTATGTACTTTCAAGCTCTGGGT 211  
QY 195 TGCCGAGCACACACTCCAAACAGAGGTGCTGACAGAGCTGCTCGTCACTCTTTGGC 254  
Db 212 TGCCGAGCACACACTCCAAACAGAGGTGCTGACAGAGCTGCTCGTCACTCTTTGGC 271  
QY 255 ACCAAGCCTCATGTGCTGCCAGCTGTCTTGGCAGAGGTGGCTCTCTGTGTACCACTGAA 314  
Db 272 ACCAAGCCTCATGTGCTGCCAGCTGTCTTGGCAGAGGTGGCTCTCTGTGTACCACTGAA 331  
QY 315 CACATCTTGGTTCGCTGGCTGCTGGGGTGTCTGTAGCACCTGGAGAGCTGCTGCC 374  
Db 332 CACATCTTGGTTCGCTGGCTGCTGGGGTGTCTGTAGCACCTGGAGAGCTGCTGCC 391  
QY 375 CCAACACACAGGTGCTGGGGTCTTGGCCAACTGCTGTGTGTCCAGGAAGGGGC 434  
Db 392 CCAACACACAGGTGCTGGGGTCTTGGCCAACTGCTGTGTGTCCAGGAAGGGGC 451  
QY 435 ATAGTAGTGGCGGGGCCGCCACGTGGGGAGCAGCAGACCAAGCTCTCTGCAGCATCTG 494  
Db 452 ATAGTAGTGGCGGGGCCGCCACGTGGGGAGCAGCAGACCAAGCTCTCTGCAGCATCTG 511  
QY 495 CTGGAGGCTGTGGGGTGTGAGAGGTGCTGAAAGCTAGCTGACATCCACACTGTC 554  
Db 512 CTGGAGGCTGTGGGGTGTGAGAGGTGCTGAAAGCTAGCTGACATCCACACTGTC 571  
QY 555 CTCAGTGGCAGTGGCTGGCTTCGTGTGTGATTCCTCGAGGCCCTGGCTCAAGAGGC 614  
Db 572 CTCAGTGGCAGTGGCTGGCTTCGTGTGTGATTCCTCGAGGCCCTGGCTCAAGAGGC 631  
QY 615 GTCAAGATGGGATGCCAGCAGCTTGGCCACCCATCGCTGCCAGACCTGCTGGGG 674  
Db 632 GTCAAGATGGGATGCCAGCAGCTTGGCCACCCATCGCTGCCAGACCTGCTGGGG 691  
QY 675 ACGGCCAAGATGCTGCTGCACAGGGCCAAACCCAGGCCAGCTGGCTCAGACCTGTGC 734  
Db 692 ACGGCCAAGATGCTGCTGCACAGGGCCAAACCCAGGCCAGCTGGCTCAGACCTGTGC 751  
QY 735 ACCCGGGTGGCAGCACCATCTATGAGCTCCACGCCCTGGAGCAGGGGGCTGGAGCA 794  
Db 752 ACCCGGGTGGCAGCACCATCTATGAGCTCCACGCCCTGGAGCAGGGGGCTGGAGCA 811  
QY 795 GCCACATGAGCGCGGTGGAGGCTGCCACCTGCCGGGCCAAGAGAGCTCAGCAGAAAGTAG 854  
Db 812 GCCACATGAGCGCGGTGGAGGCTGCCACCTGCCGGGCCAAGAGAGCTCAGCAGAAAGTAG 871  
QY 855 GCTGGGCTGTGGCCATCTTTCTGCTGTGTGGCCCTGCTCCCTGTGTGTCCCTTCCC 914  
Db 872 GCTGGGCTGTGGCCATCTTTCTGCTGTGTGGCCCTGCTCCCTGTGTGTCCCTTCCC 931  
QY 915 CTGAGGACTTGGCGCTCCCTCCCTCTGATGAGGGTCTCCCTACTGCTCTTCTCCCTTTG 974  
Db 932 CTGAGGACTTGGCGCTCCCTCCCTCTGATGAGGGTCTCCCTACTGCTCTTCTTCCCTTTG 991  
QY 975 CACAGGAAATGACAGGGGCGCAGGACTTGGAGGTTCCAGAGCGGGGGAGCCCGACCA 1034  
Db 992 CACAGGAAATGACAGGGGCGCAGGACTTGGAGGTTCCAGAGCGGGGGAGCCCGACCA 1051  
QY 1035 GTGGGACACTCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 1094  
Db 1052 GTGGGACACTCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 1111

QY 1095 GCTGCAGTGGCCACCTTGTGTCACACATTGGTTCTGAGGGGCCCAAGAGA 1145  
|||||  
Db 1112 GCTGCAGTGGCCACCTTGTGTCACACATTGGTTCTGAGGGGCCCAAGAAA 1162  
|||||  
RESULT 7  
ID ABA00763  
AC ABA00763; DNA; 860 BP.  
XX  
XX ABA00763;  
DT 18-MAR-2003 (first entry)  
DE  
XX P5CR related nucleic acid #6.  
XX  
XX Pyroline 5 carboxylate reductase; P5CR; NADPH; cancer; breast; colon;  
KW Pyroline 5 carboxylate; P5C; proline; p53; kidney; lung; ovary; ss.  
XX  
XX Homo sapiens.  
OS  
PN WO200299043-A2.  
XX  
XX 12-DEC-2002.  
XX  
XX 03-JUN-2002; 2002WO-US17319.  
XX  
XX 05-JUN-2001; 2001US-296080P.  
PR 10-OCT-2001; 2001US-328509P.  
XX  
XX (EXEL-) EXELIS INC.  
XX  
XX Friedman L, Plowman GD, Belvin M, Francis-Lang H, Li D, Funke RP;  
PI Engst S;  
PI  
XX  
XX WPI; 2003-140606/13.  
XX  
XX Identifying p53 pathway modulating agents with p5CR genes, useful for  
PT the diagnosis and treatment of disorders associated with defects in the  
PT p53 pathway, such as cancer of the breast, colon, kidneys, lung and  
PT ovary -  
XX  
XX Disclosure; Page 43-44; 53pp; English.  
XX  
XX The sequences given in ABA00758-66 represent pyrroline 5 carboxylate  
CC reductase (P5CR) related nucleic acid sequences. P5CR catalyses the  
CC NAD(P)H-dependent conversion of pyrroline 5 carboxylate (P5C) to  
CC proline. These sequences may be used in the method of the invention  
CC for identifying a candidate p53 pathway modulating agent. The method  
CC comprises providing an assay system comprising a purified P5CR  
CC polypeptide or nucleic acid, or a functionally active fragment or  
CC derivative, contacting the assay system with a test agent, where the  
CC system provides a reference activity, and detecting a test agent-biased  
CC activity of the assay system. The method of the present invention is  
CC useful for the diagnosis and treatment of disorders associated with  
CC defects in the p53 pathway, such as cancer of the breast, colon,  
CC kidneys, lung and ovary.  
XX  
SQ Sequence 860 BP; 153 A; 274 C; 291 G; 142 T; 0 other;  
  
Query Match 64.1%; Score 857; DB 25; Length 860;  
Best Local Similarity 100.0%; Pred. No. 7.5e-182;  
Matches 857; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 15 GTCCGAGGCAACAGATGGCAGTGGGAGCGCTCTCCCGCGGGTGGCTTCGTGGGC 74  
|||||  
Db 4 GTCCGAGGCAACAGATGGCAGTGGGAGCGCTCTCCCGCGGGTGGCTTCGTGGGC 63  
|||||  
QY 75 GCGGGCGCATGCGGGGGCCATFCGCGAGGCGCTCATCAGAGCAAGAAAGTGAAGCT 134  
|||||  
Db 64 GCGGGCGCATGCGGGGGCCATFCGCGAGGCGCTCATCAGAGCAAGAAAGTGAAGCT 123  
|||||  
QY 135 CACGACACTGCGCCAGTCACCAACAGACAGGACCTATGTCACCTTCAAGCTGGGT 194  
|||||

Db 124 CAGCACATACTGGCCAGTGCACCACAGACAGAAACCTATGTACCTTTCAAGCTCTGGGT 183  
|||||  
QY 195 TGGCGGACCAGCAGCTCCAAACAGAGAGTGTCTGCAGAGTGTGCTGCTCATCTTTTGGC 254  
|||||  
Db 184 TGGCGGACCAGCAGCTCCAAACAGAGAGTGTCTGCAGAGTGCCTGCTCGTCATCTTTGCC 243  
|||||  
QY 255 ACCAAGCCTCATGTGTGTCAGCTGTCTGGCAGAGTGGCTCTGTGTGTCACCACTGAA 314  
|||||  
Db 244 ACCAAGCCTCATGTGTGTCAGCTGTCTGGCAGAGTGGCTCTGTGTGTCACCACTGAA 303  
|||||  
QY 315 CACATCTTGTGTCCTGGTGTCTGAGCAGCTGTCTGAGCAGCCTTGAGGAGTGTGCTGCC 374  
|||||  
Db 304 CACATCTTGTGTCCTGGTGTCTGAGCAGCCTTGAGGAGTGTGCTGCC 363  
|||||  
QY 375 CCAACACAGGGTGTGTCGGGTCTTGGCCAACTGCTGTGTGTCAGGAGGGGCC 434  
|||||  
Db 364 CCAACACAGGGTGTGTCGGGTCTTGGCCAACTGCTGTGTGTCAGGAGGGGCC 423  
|||||  
QY 435 ATAGTGATGGCGGGGCCCGCAGCTGGGGAGGAGCAGCAAGCTTCCTGACAGCATCTG 494  
|||||  
Db 424 ATAGTGATGGCGGGGCCCGCAGCTGGGGAGGAGTGCCTGAGAGCTTCTGACATCCATCTG 483  
|||||  
QY 495 CTGAGGCGCTGTGGCGGTGTGAGGAGTGCCTGAAGCCTACGTGACATCCACACTGGC 554  
|||||  
Db 484 CTGAGGCGCTGTGGCGGTGTGAGGAGTGCCTGAGAGCTTCTGACATCCACACTGGC 543  
|||||  
QY 555 CTCAGTGGCAGTGGCGCTTCTGCTGTGTGCTATCTCCGAGGCGCTGGCTGAAGAGCC 614  
|||||  
Db 544 CTCAGTGGCAGTGGCGTGGCTTCTGCTGTGTGCTATCTCCGAGGCGCTGGCTGAAGAGCC 603  
|||||  
QY 615 GTCAGATGGCATGCCAGCAGCTGGCCACCGCATGCTGCCAGACCCCTGCTGGGG 674  
|||||  
Db 604 GTCAGATGGCATGCCAGCAGCTGGCCACCGCATGCTGCCAGACCCCTGCTGGGG 663  
|||||  
QY 675 ACGCCAAGATGCTGTGCAGAGGCGCAACACCCAGCCAGCTGGCTCAGACGTGTGC 734  
|||||  
Db 664 ACGCCCAAGATGCTGTGCAGAGGCGCAACACCCAGCCAGCTGGCTCAGACGTGTGC 723  
|||||  
QY 735 ACCCGGGTGGCACCACTATATGAGTCCACGCTCCAGCAGGCGGGCTGCGAGCA 794  
|||||  
Db 724 ACCCGGGTGGCACCACTATATGAGTCCACGCTCCAGCAGGCGGGCTGCGAGCA 783  
|||||  
QY 795 GCCACCATGAGCGCGTGGAGGCTGCCACCTGCGGGCCCAAGGCTCAGCAGAAAGTAG 854  
|||||  
Db 784 GCCACCATGAGCGCGTGGAGGCTGCCACCTGCGGGCCCAAGGCTCAGCAGAAAGTAG 843  
|||||  
QY 855 GCTGGGCTCTGGCCATC 871  
|||||  
Db 844 GCTGGGCTCTGGCCATC 860  
|||||  
RESULT 8  
AAH08500  
ID AAH08500 standard; cDNA; 794 BP.  
XX AAH08500;  
AC AAH08500;  
XX  
XX 26-JUN-2001 (first entry)  
XX  
XX Human cDNA clone (5'-primer) SEQ ID NO:5335.  
DE  
XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.  
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.  
XX  
XX Homo sapiens.  
OS  
XX EP1074617-A2.  
PN  
XX 07-FEB-2001.  
PD  
XX 28-JUL-2000; 2000EP-0116126.  
PF  
XX 29-JUL-1999; 99JP-0248036.  
PR 27-AUG-1999; 99JP-0300253.  
PR







Db 3795 GNCCTGAGCAG-CACCATGAGCGCGTGGAGG-TGCCACCTGCGGGCCAAAGAGCTCA 3852  
QY 844 GCAGAAAGTAGCTGGGCTCTGGCCATCCTTTCTGCTCTCTGCTGCTGCTCTCCCTG 903  
Db 3853 GCAGAAAGTAGCTGGGCTCTGGG-CATCCTTTCTGCTCTCTGCTGCTCTCCCTG 3911  
QY 904 TCTCCTCTCCCTGAGGACTGCGGCTCCCTCCCTCCTCATGAGGCTCTCCTACTGCTCC 963  
Db 3912 TGTCTCTCTCCCTGAGGACTGCGGCTCCCTCCCTCCTCATGAGGCTCTCCTACTGCTCC 3971  
QY 964 TTCTCTCCCTTGACAGGAAATGACAGGGGAGGACTTTGGAGGTTTCAGCAGCGGGGG 1023  
Db 3972 TTCTCTCCCTTGACAGGAAATGACAGGGGAGGACTTTGGAGGTTTCAGCAGCGGGGG 4031  
QY 1024 AGCCCGGACCACTGAGGACCTCCCTCCCTCCCTGAGCAGAGGACCCCTGCTGTGG 1083  
Db 4032 AG-CCCGACCACTGAGGACCTCCCTCCCTCCCTGAGCAGAGGACCCCTGCTGTGG 4090  
QY 1084 CTCTGCCCCCTTGCAGTGAGCCACCTTTGTCGCAACATTTGGTTCTGAGGGGCCAAGA 1143  
Db 4091 CTCTGCCCCCTTGCAGTGAGCCACCTTTGTCGCAACATTTGGTTCTGAGGGGCCAAGA 4150  
QY 1144 GATGGCGCTTGGTCATTTGCCCGCATGTTGGGCGAGTTGGTTGAGGCCATGAACAGAAC 1203  
Db 4151 GATGGCGCTTGGTCATTTGCCCGCATGTTGGGCGAGTTGGTTGAGGCCATGAACAGAAC 4210  
QY 1204 TTACGGTTACAGGACGCGCTGSCCAATGCTGGTCTGAGCTGGAGCTTGCCTTTGGCT 1263  
Db 4211 TTACGGTTACAGGACGCGCTGSCCAATGCTGGTCTGAGCTGGAGCTTGCCTTTGGCT 4270  
QY 1264 TTCCAAGTGGGCTGTCAGCTACAGCCAGGCGGCTGCCTCATCTCAGCTCTAGGGGGC 1323  
Db 4271 TTCCAGGTGGCTCCCTGTCAGCTACAGCCAGGCGGCTGCCTCATCTCAGCTCTAGGGGGC 4330  
QY 1324 ACGACCATATGGG 1337  
Db 4331 ACGACCATATGGG 4344

RESULT 11  
AAH06264  
ID AAH06264 standard; cDNA; 563 BP.  
XX  
AC AAH06264;  
XX  
DT 26-JUN-2001 (first entry)  
XX  
DE Human cDNA clone (5'-primer) SEQ ID NO:3099.  
XX  
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.  
XX  
OS Homo sapiens.  
XX  
PN EP1074617-A2.  
XX  
PD 07-FEB-2001.  
XX  
PF 28-JUL-2000; 2000EP-0116126.  
XX  
PR 29-JUL-1999; 99JP-0248036.  
PR 27-AUG-1999; 99JP-0300253.  
PR 11-JAN-2000; 2000JP-0118776.  
PR 02-MAY-2000; 2000JP-0183767.  
PR 09-JUN-2000; 2000JP-0241899.  
XX  
PA (HELI-) HELIX RES INST.  
XX  
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
XX  
DR WPI; 2001-318749/34.  
XX

PT Primer sets for synthesizing polynucleotides, particularly the 5602  
PT full-length cDNAs defined in the specification, and for the detection  
PT and/or diagnosis of the abnormality of the proteins encoded by the  
PT full-length cDNAs -  
XX Claim 1; SEQ ID 3099; 2537pp + CD ROM; English.  
XX  
CC The present invention describes primer sets for synthesizing 5602  
CC full-length cDNAs defined in the specification. Where a primer set  
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary  
CC to the complementary strand of a polynucleotide which comprises one of  
CC the 5602 nucleotide sequences defined in the specification, where the  
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
CC of an oligonucleotide comprising a sequence complementary to the  
CC complementary strand of a polynucleotide which comprises a 5'-end  
CC sequence and an oligonucleotide comprising a sequence complementary to a  
CC polynucleotide which comprises a 3'-end sequence, where the  
CC oligonucleotide comprises at least 15 nucleotides and the combination of  
CC the 5'-end sequence/3'-end sequence is selected from those defined in  
CC the specification. The primer sets can be used in antisense therapy and  
CC in gene therapy. The primers are useful for synthesizing polynucleotides,  
CC particularly full-length cDNAs. The primers are also useful for the  
CC detection and/or diagnosis of the abnormality of the proteins encoded by  
CC the full-length cDNAs. The primers allow obtaining of the full-length  
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to  
CC AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632  
CC represent oligonucleotides, all of which are used in the exemplification  
CC of the present invention.  
XX  
SQ Sequence 563 BP; 99 A; 171 C; 181 G; 108 T; 4 other;

Query Match 36.4%; Score 486.6; DB 22; Length 563;  
Best Local Similarity 95.9%; Pred. No. 3.3e-99;  
Matches 540; Conservative 0; Mismatches 18; Indels 5; Gaps 4;  
QY 27 AAGATGGCAGCTGC-GGAGCCGCTCTCCGGCGCGCTGGGCTTCGTTGGCGCGCGCGCAT 85  
Db 1 AAGATGGCAGCTGGGGAGCGGCTCTCCGGCGCGCTGGGCTTCGTTGGCGCGCGCGCAT 60  
QY 86 GCGGGGGGCCATCGCGCAGGGCCCTCATCAGAGCAGGAAAGTGAAGCTCAGACATACT 145  
Db 61 GCGGGGGGCCATCGCGCAGGGCCCTCATCAGAGCAGGAAAGTGAAGCTCAGACATACT 120  
QY 146 GCGCAGTGCACCAACAGACAGAACCTATGTCTACTTTCAAGCTCTGGTTGCGGACCAC 205  
Db 121 GCGCAGTGCACCAACAGACAGAACCTATGTCTACTTTCAAGCTCTGGTTGCGGACCAC 180  
QY 206 GCACTCCAACAGGAGGTGCTGCAGAGCTGCCTGCTGCTCATCTTTGCCACCAAGCCTCA 265  
Db 181 GCACTCCAACAGGAGGTGCTGCAGAGCTGCCTGCTGCTCATCTTTGCCACCAAGCCTCA 240  
QY 266 TGTGCTGCCAGCTGTCTCTGG-CAGAGGTGCTCTCTGTGTGTCACCTGAACATCTTGG 324  
Db 241 TGTGCTGCCAGCTGTCTCTGGCAGAGGTGGCTCTCTGTGTGTCACCTGAACATCTTGG 300  
QY 325 TGTCCGTGGCTGTGGGGTGTCTCTGAGCACCCTGAGAGGAGCTGCTGCCCCCAACACAC 384  
Db 301 TGTCCGTGGCTGTGGGGTGTCTCTGAGCACCCTGAGAGGAGCTGCTGCCCCCAACACAC 360  
QY 385 GGTGCTGCGGGTCTTGGCCCAACTGCTGCTGGTCCAGAGGAGGCGCATAG--TGAT 442  
Db 361 GGTGCTGCGGGTCTTGGCCCAACTGCTGCTGGTCCAGAGGAGGCGCATAGATNATG 420  
QY 443 GCGCGGGGGCGCCACGCTGGGGAGCAGCAGCAAGCTCCTGCAGCATCTGCTGGAGGC 502  
Db 421 GCCTCGGGGGCGCCACGCTGGGAGCAGCAGCAAGCTCCTGCAGCATCTTCTGGATGC 480  
QY 503 CTGTGGCGGTGTGAGGAGGTGCTGAAGCTTACGCTGACATCCACACTGGCTCAGTGG 562  
Db 481 CTGTGGCGGTGTGAGGAGGTGCTGAATCTTACGCTGACATCCACACTGGCTCAGTGG 540  
QY 563 CAGTGG-CGTGGCTTCGTCTGT 584

```
DB      541  CAGTGGTCGTNGCCTCCNGTNT 563
RESULT 12
AAS41342
ID      AAS41342 standard; cDNA; 270 BP.
XX      AC      AAS41342;
XX      XX      17-DEC-2001 (first entry)
XX      DE      cDNA encoding novel human enzyme polypeptide #558.
XX      KW      Human; oxidoreductase enzyme; transferase; hydrolase; lyase; isomerase;
XX      KW      ligase; hyperproliferative disorder; immunodeficiency disorder;
XX      KW      autoimmune disorder; neurological disorder; metabolic disorder;
XX      KW      inflammatory disorder; cardiovascular disorder; reproductive disorder;
XX      KW      blood-related disorder; infectious disorder; gene therapy; cytostatic;
XX      KW      anti arthritic; nephrotropic; anticoagulant; ss.
XX      OS      Homo sapiens.
XX      XX      PN      W0200155301-A2.
XX      XX      PD      02-AUG-2001.
XX      XX      PF      17-JAN-2001; 2001WO-US01239.
XX      XX      PR      31-JAN-2000; 2000US-0179065.
XX      PR      04-FEB-2000; 2000US-0180628.
XX      PR      24-FEB-2000; 2000US-0184664.
XX      PR      02-MAR-2000; 2000US-0186350.
XX      PR      16-MAR-2000; 2000US-0189874.
XX      PR      17-MAR-2000; 2000US-0190076.
XX      PR      18-APR-2000; 2000US-0198123.
XX      PR      19-MAY-2000; 2000US-0205515.
XX      PR      07-JUN-2000; 2000US-0209467.
XX      PR      28-JUN-2000; 2000US-0214886.
XX      PR      30-JUN-2000; 2000US-0215135.
XX      PR      07-JUL-2000; 2000US-0216647.
XX      PR      11-JUL-2000; 2000US-0216880.
XX      PR      11-JUL-2000; 2000US-0217487.
XX      PR      11-JUL-2000; 2000US-0217496.
XX      PR      14-JUL-2000; 2000US-0218290.
XX      PR      26-JUL-2000; 2000US-0220963.
XX      PR      14-AUG-2000; 2000US-0224518.
XX      PR      14-AUG-2000; 2000US-0224519.
XX      PR      14-AUG-2000; 2000US-0225213.
XX      PR      14-AUG-2000; 2000US-0225214.
XX      PR      14-AUG-2000; 2000US-0225266.
XX      PR      14-AUG-2000; 2000US-0225267.
XX      PR      14-AUG-2000; 2000US-0225268.
XX      PR      14-AUG-2000; 2000US-0225270.
XX      PR      14-AUG-2000; 2000US-0225447.
XX      PR      14-AUG-2000; 2000US-0225757.
XX      PR      14-AUG-2000; 2000US-0225758.
XX      PR      18-AUG-2000; 2000US-0225759.
XX      PR      22-AUG-2000; 2000US-0226681.
XX      PR      22-AUG-2000; 2000US-0226686.
XX      PR      22-AUG-2000; 2000US-0227182.
XX      PR      23-AUG-2000; 2000US-0227009.
XX      PR      30-AUG-2000; 2000US-0228924.
XX      PR      01-SEP-2000; 2000US-0229287.
XX      PR      01-SEP-2000; 2000US-0229343.
XX      PR      01-SEP-2000; 2000US-0229344.
XX      PR      05-SEP-2000; 2000US-0229345.
XX      PR      05-SEP-2000; 2000US-0229509.
XX      PR      06-SEP-2000; 2000US-0230437.
XX      PR      06-SEP-2000; 2000US-0230438.
XX      PR      08-SEP-2000; 2000US-0231242.
XX      PR      08-SEP-2000; 2000US-0231243.
XX      PR      08-SEP-2000; 2000US-0231244.
XX      PR      08-SEP-2000; 2000US-0231413.
XX      PR      08-SEP-2000; 2000US-0231414.
XX      PR      08-SEP-2000; 2000US-0232080.
XX      PR      08-SEP-2000; 2000US-0232081.
XX      PR      12-SEP-2000; 2000US-0231968.
XX      PR      14-SEP-2000; 2000US-0232397.
XX      PR      14-SEP-2000; 2000US-0232398.
XX      PR      14-SEP-2000; 2000US-0232399.
XX      PR      14-SEP-2000; 2000US-0232400.
XX      PR      14-SEP-2000; 2000US-0232401.
XX      PR      14-SEP-2000; 2000US-0233063.
XX      PR      14-SEP-2000; 2000US-0233064.
XX      PR      14-SEP-2000; 2000US-0233065.
XX      PR      21-SEP-2000; 2000US-0234223.
XX      PR      21-SEP-2000; 2000US-0234274.
XX      PR      25-SEP-2000; 2000US-0234997.
XX      PR      25-SEP-2000; 2000US-0234998.
XX      PR      26-SEP-2000; 2000US-0235484.
XX      PR      27-SEP-2000; 2000US-0235834.
XX      PR      27-SEP-2000; 2000US-0235836.
XX      PR      29-SEP-2000; 2000US-0236327.
XX      PR      29-SEP-2000; 2000US-0236367.
XX      PR      29-SEP-2000; 2000US-0236368.
XX      PR      29-SEP-2000; 2000US-0236369.
XX      PR      29-SEP-2000; 2000US-0236370.
XX      PR      02-OCT-2000; 2000US-0236802.
XX      PR      02-OCT-2000; 2000US-0237037.
XX      PR      02-OCT-2000; 2000US-0237038.
XX      PR      02-OCT-2000; 2000US-0237039.
XX      PR      13-OCT-2000; 2000US-0237040.
XX      PR      13-OCT-2000; 2000US-0239935.
XX      PR      13-OCT-2000; 2000US-0239937.
XX      PR      20-OCT-2000; 2000US-0240960.
XX      PR      20-OCT-2000; 2000US-0241221.
XX      PR      20-OCT-2000; 2000US-0241785.
XX      PR      20-OCT-2000; 2000US-0241786.
XX      PR      20-OCT-2000; 2000US-0241808.
XX      PR      20-OCT-2000; 2000US-0241809.
XX      PR      20-OCT-2000; 2000US-0241826.
XX      PR      01-NOV-2000; 2000US-0244617.
XX      PR      08-NOV-2000; 2000US-0246474.
XX      PR      08-NOV-2000; 2000US-0246475.
XX      PR      08-NOV-2000; 2000US-0246476.
XX      PR      08-NOV-2000; 2000US-0246477.
XX      PR      08-NOV-2000; 2000US-0246478.
XX      PR      08-NOV-2000; 2000US-0246523.
XX      PR      08-NOV-2000; 2000US-0246524.
XX      PR      08-NOV-2000; 2000US-0246525.
XX      PR      08-NOV-2000; 2000US-0246526.
XX      PR      08-NOV-2000; 2000US-0246527.
XX      PR      08-NOV-2000; 2000US-0246528.
XX      PR      08-NOV-2000; 2000US-0246532.
XX      PR      08-NOV-2000; 2000US-0246609.
XX      PR      08-NOV-2000; 2000US-0246610.
XX      PR      08-NOV-2000; 2000US-0246611.
XX      PR      08-NOV-2000; 2000US-0246613.
XX      PR      17-NOV-2000; 2000US-0249207.
XX      PR      17-NOV-2000; 2000US-0249208.
XX      PR      17-NOV-2000; 2000US-0249209.
XX      PR      17-NOV-2000; 2000US-0249210.
XX      PR      17-NOV-2000; 2000US-0249211.
XX      PR      17-NOV-2000; 2000US-0249212.
XX      PR      17-NOV-2000; 2000US-0249213.
XX      PR      17-NOV-2000; 2000US-0249214.
XX      PR      17-NOV-2000; 2000US-0249215.
XX      PR      17-NOV-2000; 2000US-0249216.
XX      PR      17-NOV-2000; 2000US-0249217.
XX      PR      17-NOV-2000; 2000US-0249218.
XX      PR      17-NOV-2000; 2000US-0249244.
```

```
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX
XX Rosen CA, Barash SC, Ruben SM;
PI
XX
XX WPI: 2001-465566/50.
DR P-PSDB; AAU23472.
XX
XX Novel polypeptides and polynucleotides useful for diagnosing,
PT preventing, treating neural, immune system, muscular, reproductive,
PT pulmonary, cardiovascular, renal, proliferative disorders and cancerous
PT diseases
XX
XX Claim 4; SEQ ID NO 568; 1180pp; English.
PS
XX
XX The present invention relates to the isolation of novel human enzyme
CC polypeptides (AAU22915-AAU23814), and the cDNA and genomic sequences
CC encoding them. The enzyme polypeptides of the invention may comprise the
CC functional classes of oxidoreductases, transferases, hydrolases, lyases,
CC isomerases or ligases. The sequences of the invention are useful in the
CC diagnosis, treatment, prevention and/or prognosis of a wide range of
CC disorders, including hyperproliferative disorders (e.g. cancer),
CC immunodeficiency disorders (e.g. AIDS) autoimmune disorders
CC (e.g. arthritis), neurological disorders (e.g. Alzheimer's disease),
CC metabolic disorders (e.g. phenylketonuria), inflammatory disorders
CC (e.g. asthma), cardiovascular disorders (e.g. atherosclerosis),
CC blood-related disorders (e.g. haemophilia), reproductive disorders
CC (e.g. infertility) and infectious disorders (e.g. influenza). The
CC polynucleotides of the invention can also be used in gene therapy.
CC AAU40785-AAU41684 represent cDNA sequences encoding for the novel human
CC enzyme polypeptides of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 270 BP; 54 A; 82 G; 45 T; 1 other;
SQ
Query Match 18.1%; Score 241.8; DB 22; Length 270;
Best Local Similarity 95.9%; Pred. No. 1.3e-44;
Matches 258; Conservative 1; Mismatches 8; Indels 2; Gaps 1;
QY 10 AGTCTTCGAGGCACACAGATGGCAGCTGGCGAGCCGCTCTCCGGCGCGCTGGGCTTCG 69
DB 11 |||||||
DB 2 AGGCGCTCCGAGGACAAAGATGGCAGCTGGCGAGCCGCTCTCCGGCGCGCTGGGCTTCG 61
QY 70 TGGCGCGCGGCGCGCATGGCGGGGCCATCGCGAGGGCCCTCATCAGACAGGAAAGTGG 129
DB 62 TGGCGCGGCGCGCATGGCGGGGCCATCGCGAGGGCCCTCATCAGACAGGAAAGTGG 121
QY 130 AAGCTCAGCACATACCTGGCGAGTGGCCACCAACAGACAGGAAACCTATGTCTACTTTCAGGCTC 189
DB 122 AAGCTCAGCACATACCTGGCGAGTGGCCACCAACAGACAGGAAACCTATGTCTACTTTCAGGCTC 181
QY 190 TGGGTTGCCGGACCGACACTCCACACGAGGAGGTGCTGCAGAGCT--GCCTGCTCGTCAT 247
|||||
182 TGGGTTGCCGGACCGACACTCCACACGAGGAGGTGCTGCAGAGCTGCTGCTCAT 241
|||||
248 CTTTGGCACCAGCCTCATGTGTGTCAG 276
|||||
242 CTTTGGCACCAGCCTCATGTGTGTCAG 270
RESULT 13
ABA00760
ID ABA00760 standard; DNA; 1848 BP.
XX
XX ABA00760;
XX
XX 18-MAR-2003 (first entry)
XX
XX p5CR related nucleic acid #3.
XX
XX Pyrroline 5 carboxylate reductase; p5CR; NADPH; cancer; breast; colon;
KW pyrroline 5 carboxylate; p5C; proline; p53; kidney; lung; ovary; ss.
XX
XX Homo sapiens.
XX
XX WO200299043-A2.
XX
XX 12-DEC-2002.
XX
XX 03-JUN-2002; 2002WO-US17319.
XX
XX 05-JUN-2001; 2001US-296080P.
PR 10-OCT-2001; 2001US-328509P.
XX
XX (EXEL-) EXELIS INC.
XX
XX Friedman L, Plovman GD, Belvin M, Francis-Lang H, Li D, Funke RP;
PI Engst S;
XX
XX WPI: 2003-140606/13.
XX
XX Identifying p53 pathway modulating agents with p5CR genes, useful for
PT the diagnosis and treatment of disorders associated with defects in the
PT p53 pathway, such as cancer of the breast, colon, kidneys, lung and
PT ovary
XX
XX Disclosure; Page 40-41; 53pp; English.
XX
XX The sequences given in ABA00758-66 represent pyrroline 5 carboxylate
CC reductase (P5CR) related nucleic acid sequences. P5CR catalyses the
CC NAD(P)H-dependent conversion of pyrroline 5 carboxylate (P5C) to
CC proline. These sequences may be used in the method of the invention
CC for identifying a candidate p53 pathway modulating agent. The method
CC comprises providing an assay system comprising a purified P5CR
CC polypeptide or nucleic acid, or a functionally active fragment or
CC derivative, contacting the assay system with a test agent, where the
CC system provides a reference activity, and detecting a test agent-biased
CC activity of the assay system. The method of the present invention is
CC useful for the diagnosis and treatment of disorders associated with
CC defects in the p53 pathway, such as cancer of the breast, colon,
CC kidneys, lung and ovary.
XX
XX Sequence 1848 BP; 364 A; 582 C; 541 G; 361 T; 0 other;
SQ
Query Match 16.9%; Score 226.4; DB 25; Length 1848;
Best Local Similarity 58.1%; Pred. No. 5.3e-41;
Matches 467; Conservative 0; Mismatches 316; Indels 21; Gaps 3;
QY 56 GCGCGTGGGCTCTGTGGCGGGCGCGATGCGGGGGCCATCGCGAGGCGCTTCATCAG 115
DB 52 GAGCGTGGGCTCTATCGCGCTGGCGCTGGCGCTGGCTTTGGCTTGGCCAAAGGCTTCACAG 111
QY 116 AGCAGAAAGTGAAGCTCAGCACATCTGGCCAGTGGCCAAACAGACAGGAACTATG 175
DB 112 AGCAGGCGTCTTGGCTGCCCAAGATAATGGCTAGCTCCC---CAGACATGGACCTGGC 168
|||||
```





